

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 1, 2004, 16:15:10 ; Search time 2064 Seconds
(without alignments)
2823.899 Million cell updates/sec

Title: US-09-993-808B-1
Perfect score: 1372
Sequence: 1 cccacggtccggacgcaag.....agtgacacctaattcattc 1372

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|--------|---------------|--------|-------|-------------|
| 1 | 1371.6 | 100.0 | 1372 | 7 | ABV74603 |
| 2 | 644.2 | 47.0 | 723 | 3 | AA95281 |
| 3 | 351 | 25.6 | 841 | 7 | ABV74605 |
| 4 | 299 | 21.8 | 1242 | 6 | ABK93958 |
| 5 | 234.6 | 17.1 | 639 | 3 | AAA95280 |
| 6 | 223.2 | 16.3 | 572 | 3 | AA02390 |
| 7 | 223.2 | 16.3 | 572 | 3 | AAA95276 |
| 8 | 219 | 16.0 | 568 | 6 | ABK93954 |
| 9 | 114.8 | 8.4 | 2015 | 2 | AAZ24904 |
| 10 | 110.8 | 8.1 | 3313 | 5 | AAZ29102 |
| 11 | 110.8 | 8.1 | 3313 | 6 | ABS68242 |
| 12 | 110.8 | 8.1 | 3313 | 9 | ADC25236 |
| 13 | 110.2 | 8.0 | 1557 | 3 | AAC99034 |
| 14 | 110.2 | 8.0 | 1557 | 5 | AAZ00830 |
| 15 | 109.6 | 8.0 | 701 | 4 | AAH33453 |
| 16 | 108.8 | 7.9 | 1316 | 3 | AAC59966 |
| 17 | 108.4 | 7.9 | 149 | 4 | AA02569 |
| 18 | 108.4 | 7.9 | 2116 | 3 | AAC99033 |
| 19 | 108.4 | 7.9 | 2116 | 5 | AAZ00847 |
| 20 | 107.4 | 7.8 | 577 | 3 | AAZ78234 |
| 21 | 107.4 | 7.8 | 1837 | 3 | AAC98918 |
| 22 | 107.4 | 7.8 | 2140 | 4 | ABA06584 |
| 23 | 107.4 | 7.8 | 2140 | 6 | ABV83921 |

| | | | | | | | | | |
|---|----|-------|-----|------|---|----------|----------|-------|-------|
| C | 24 | 107.2 | 7.8 | 2322 | 7 | ABX12257 | Abx12257 | cdna | enco |
| | 25 | 106.8 | 7.8 | 845 | 6 | ABQ54253 | Abq54253 | Human | ova |
| | 26 | 106.6 | 7.8 | 858 | 4 | AAH33135 | Aah33135 | Human | col |
| | 27 | 106.4 | 7.8 | 584 | 3 | AAC98302 | Aac98302 | Human | col |
| | 28 | 106.4 | 7.8 | 603 | 5 | ABA11175 | Abal1175 | Human | ner |
| | 29 | 106.4 | 7.8 | 763 | 3 | AAC98303 | Aac98303 | Human | col |
| | 30 | 106.4 | 7.8 | 824 | 2 | AAZ37387 | Aax37387 | Human | sec |
| | 31 | 105.6 | 7.7 | 2402 | 7 | ABX12256 | Abx12256 | cdna | enco |
| | 32 | 105.4 | 7.7 | 418 | 3 | AAC98309 | Aac98309 | Human | col |
| | 33 | 105.4 | 7.7 | 627 | 4 | AA02583 | Aal02583 | Human | rep |
| | 34 | 105.4 | 7.7 | 627 | 4 | ABL97313 | Ab197313 | Human | tes |
| | 35 | 105.4 | 7.7 | 1280 | 4 | AAH34250 | Aah34250 | Human | col |
| | 36 | 105.4 | 7.7 | 1280 | 6 | ABQ54195 | Abq54195 | Human | ova |
| | 37 | 105.4 | 7.7 | 1281 | 3 | AAC66412 | Aac66412 | Human | sec |
| | 38 | 105.2 | 7.7 | 1928 | 6 | ABK69104 | Abk69104 | DNA | encod |
| | 39 | 105.2 | 7.7 | 2454 | 3 | AAC69584 | Aac69584 | Human | sec |
| | 40 | 105.2 | 7.7 | 2454 | 7 | ADA40257 | Ada40257 | Human | sec |
| | 41 | 105.2 | 7.7 | 2454 | 9 | ADC73781 | Adc73781 | Human | sec |
| | 42 | 104.4 | 7.6 | 453 | 3 | AAC98320 | Aac98320 | Human | col |
| | 43 | 104.4 | 7.6 | 540 | 3 | AAC98305 | Aac98305 | Human | col |
| | 44 | 104.4 | 7.6 | 1270 | 3 | AAC59584 | Aac59584 | Human | sec |
| | 45 | 104.4 | 7.6 | 2152 | 4 | AAD05156 | Aad05156 | Human | sec |

ALIGNMENTS

RESULT 1
ABV74603

ID ABV74603 standard; DNA; 1372 BP.

XX AC ABV74603;

XX DT 21-FEB-2003 (first entry)

XX DE Maize CKI_B coding sequence.

XX KW Maize; cyclin-dependent kinase inhibitor; CKI; CKI_B; plant; corn;
KW crop yield; root size; plant growth; tassel size; ear size;
KW male sterility; endoreduplication; gene; ds.

XX OS Zea mays.

XX FH Key Location/Qualifiers

FT CDS 134..904

FT /*tag= a

FT /product= "CKI_B"

XX PN WO200281623-A2.

XX PD 17-OCT-2002.

XX PF 06-NOV-2001; 2001WO-US044038.

XX PR 07-NOV-2000; 2000US-0246349P.

XX PA (PION-) PIONEER HI-BRED INT INC.

XX PA (ARIZ-) ARIZONA BOARD OF REGENTS.

XX PI Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;

XX XX WPI; 2003-058511/05.

XX DR P-PSDB; ABB98757.

XX PT Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
PT involved in cell cycle regulation, and useful for altering cell cycle
PT protein content, cell cycle progression, cell number and composition of
PT plants.

XX PS Claim 1; Page 64-65; 69pp; English.

XX CC The present sequence is the coding sequence (I) for maize cyclin-

XX CC dependent kinase inhibitor (CKI), CKI_B. (I) is useful for modulating the

activity of cyclin-dependent kinase (CDK) in a plant such as a corn, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley, oil-seed Brassica and millet. Modulating the activity of CDK, preferably modulating downward is useful for providing differential growth in a plant, especially a positive growth advantage and modulating CDK activity upward is useful for increasing crop yield, root size, plant growth, tassel size and/or ear size. Modulating CDK activity is also useful for conferring male sterility and for improving transformation frequencies by increasing the number of cells in cell division. CDK activity can also be modulated for modulating endoreduplication in the endosperm of corn, sorghum, wheat, rice, barley, and millet, where the promoter used is an endosperm-preferred promoter. The cell numbers are modulated in one or more tissues of a plant, comprising root, seed, tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves, stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm, pericarp, meristematic or leaf cells. (I) is also useful for identifying maize CKI interacting proteins, by adducting the nucleic acid sequence to a second nucleic acid sequence encoding a DNA-binding domain.

SQ Sequence 1372 BP; 292 A; 396 C; 442 G; 241 T; 0 U; 1 Other;

Query Match 100.0%; Score 1371.6; DB 7; Length 1372;
Best Local Similarity 100.0%; Pred. No. 1.5e-270;
Matches 1372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| QY | 1 | CCACGCGTCCGGACGCAAGCGGCTGCAGGCAGCAGCGCCGCGCAGGCGTTGTGGCCCTGT | 60 |
| Db | 1 | CCACGCGTCCGGACGCAAGCGGCTGCAGGCAGCAGCGCCGCGCAGGCGTTGTGGCCCTGT | 60 |
| QY | 61 | GGGAGAGGAAAAAGAGAAAGAGAAACCGGCCAAGACAAGCAAGCGAGAGGCCAGGGCCGC | 120 |
| Db | 61 | GGGAGAGGAAAAAGAGAAAGAGAAACCGGCCAAGACAAGCAAGCGAGAGGCCAGGGCCGC | 120 |
| QY | 121 | GGCGTTGCGTCAGATGGGAAGTACATGCGCAAGTGCAGGGGCGCCGCGAGGCGCGGAGGT | 180 |
| Db | 121 | GGCGTTGCGTCAGATGGGAAGTACATGCGCAAGTGCAGGGGCGCCGCGAGGCGCGGAGGT | 180 |
| QY | 181 | CGCGCGCGTCGAGGTTACCGCAGGTCGTGGCGTCCGGACGAGTCCAGGTCGCGGCGGC | 240 |
| Db | 181 | CGCGCGCGTCGAGGTTACCGCAGGTCGTGGCGTCCGGACGAGTCCAGGTCGCGGCGGC | 240 |
| QY | 241 | GACCGGCGGTGTCGCAAGGTCGCCCGGAGGAGAGAGGGCGCGCGCGGGGAGCCTGC | 300 |
| Db | 241 | GACCGGCGGTGTCGCAAGGTCGCCCGGAGGAGAGAGGGCGCGCGCGGGGAGCCTGC | 300 |
| QY | 301 | TGCCGCGGTGAGCGCTGGTGGGACGGCGGAAGCTGCTACATCCACCTCGTAGCCGCAT | 360 |
| Db | 301 | TGCCGCGGTGAGCGCTGGTGGGACGGCGGAAGCTGCTACATCCACCTCGTAGCCGCAT | 360 |
| QY | 361 | GCTGTTCAFGCACCGCCTCAGCCGCGAGCCGTCGGTTGACTCGGTTCCGACCCCGTGGA | 420 |
| Db | 361 | GCTGTTCAFGCACCGCCTCAGCCGCGAGCCGTCGGTTGACTCGGTTCCGACCCCGTGGA | 420 |
| QY | 421 | GGCTGCTGATGGCGCTGCAGGACAGCGGCGCGCTCGCGGCGCGGCTCTCGCGTTG | 480 |
| Db | 421 | GGCTGCTGATGGCGCTGCAGGACAGCGGCGCGCTCGCGGCGCGGCTCTCGCGTTG | 480 |
| QY | 481 | CTCAGACCGCGCTCGTCGGTGAATTGGGCTTGGGSGTTCAGCGGGGAGCCACACCTG | 540 |
| Db | 481 | CTCAGACCGCGCTCGTCGGTGAATTGGGCTTGGGSGTTCAGCGGGGAGCCACACCTG | 540 |
| QY | 541 | CCGTCCTACGACGCTGCAGAGGCTGGCGGGATCAGTCTGTTGATGTCTCGGCGGC | 600 |
| Db | 541 | CCGTCCTACGACGCTGCAGAGGCTGGCGGGATCAGTCTGTTGATGTCTCGGCGGC | 600 |
| QY | 601 | GAGCAACTCCGGAGCGGCCACAGCCCGAGAGGCGAGAGACGACGCCATCGAGCCGGGC | 660 |
| Db | 601 | GAGCAACTCCGGAGCGGCCACAGCCCGAGAGGCGAGAGACGACGCCATCGAGCCGGGC | 660 |
| QY | 661 | GCACGGCAGCTCAGCGATCTGGAGTCGGATCTGGCGGGGCAACAAGACTGGCCCGTCGCT | 720 |
| Db | 661 | GCACGGCAGCTCAGCGATCTGGAGTCGGATCTGGCGGGGCAACAAGACTGGCCCGTCGCT | 720 |

| | | | |
|----|------|---|------|
| QY | 721 | ACCGGGCAACGCGGCTGCGAGCTGATCGTGCGCCAGCACACGAGATCCAGGAGTT | 780 |
| Db | 721 | ACCGGGCAACGCGGCTGCGAGCTGATCGTGCGCCAGCACACGAGATCCAGGAGTT | 780 |
| QY | 781 | CTTCGCGCGCGCGAGCGGCCAGGCCAAGCGCTTTGCTTCCAAGTACAACTTCGACTT | 840 |
| Db | 781 | CTTCGCGCGCGCGAGCGGCCAGGCCAAGCGCTTTGCTTCCAAGTACAACTTCGACTT | 840 |
| QY | 841 | CGTCCGCGCGCTGCCCTCGACGCGCGCGCGGTCGAGTGGCGCGCGTGGTGGTTCAGCAT | 900 |
| Db | 841 | CGTCCGCGCGCTGCCCTCGACGCGCGCGCGGTCGAGTGGCGCGCGTGGTGGTTCAGCAT | 900 |
| QY | 901 | CTGAAGCGAGCTGCGTCCGGTGCAAGTGAAGCTAGAAAGAGAAAAGATGCCCCCCC | 960 |
| Db | 901 | CTGAAGCGAGCTGCGTCCGGTGCAAGTGAAGCTAGAAAGAGAAAAGATGCCCCCCC | 960 |
| QY | 961 | CCCCCCCCCAACAACATAACGGAGAGAGAAAAACCAACAATTAAACAGACTTTATATA | 1020 |
| Db | 961 | CCCCCCCCCAACAACATAACGGAGAGAGAAAAACCAACAATTAAACAGACTTTATATA | 1020 |
| QY | 1021 | GCCTAAGCTAACCAACCACCATTCATCTCGTCCAAATGCATGCCCTTGCTTTCTCTGGAGC | 1080 |
| Db | 1021 | GCCTAAGCTAACCAACCACCATTCATCTCGTCCAAATGCATGCCCTTGCTTTCTCTGGAGC | 1080 |
| QY | 1081 | TAGCAGGAGCGTAGTTATTATTTAGTACTACTTTACTTTATTCAGAGTTATCTTGACCCC | 1140 |
| Db | 1081 | TAGCAGGAGCGTAGTTATTATTTAGTACTACTTTACTTTATTCAGAGTTATCTTGACCCC | 1140 |
| QY | 1141 | GATAGATCAATCCGCTTACTGTGTAATTTCTCTCATGCATCTCTTAGATGGAGTTTAATC | 1200 |
| Db | 1141 | GATAGATCAATCCGCTTACTGTGTAATTTCTCTCATGCATCTCTTAGATGGAGTTTAATC | 1200 |
| QY | 1201 | GTCTTAATTTATTACTGTACAGCAGCTTGSTTGGCTTGCAAGAAAAGATCTGGTTTGCT | 1260 |
| Db | 1201 | GTCTTAATTTATTACTGTACAGCAGCTTGSTTGGCTTGCAAGAAAAGATCTGGTTTGCT | 1260 |
| QY | 1261 | CAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACG | 1320 |
| Db | 1261 | CAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACG | 1320 |
| QY | 1321 | TACGCGTCATCGCAGCTCATAGCTCTTCTATAGTGTCAACCTAATTCATTC | 1372 |
| Db | 1321 | TACGCGTCATCGCAGCTCATAGCTCTTCTATAGTGTCAACCTAATTCATTC | 1372 |

RESULT 2

AAA95281

ID AAA95281 standard; cDNA; 723 BP.

AC AAA95281:

XX

DT 17-JAN-2001 (first entry)

DE Corn cyclin-dependent kinase inhibitor coding sequence #3.

KW Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
KW cell growth; herbicide; ss.

OS Zea mays.

| FH | Key | Location/Qualifiers |
|----|-----|---------------------|
|----|-----|---------------------|

| | | |
|----|-----|-------|
| FT | CDS | 3.380 |
|----|-----|-------|

```
ET
/*tag= a
```

```
FT
/product= "CDKI"
```

FT /partial

PN WO200060087-A2.

PD 12-OCT-2000.



PF 06-APR-2000; 2000WO-US009106.

XX

XX
PD 13-MAR-2003.
XX
PF 07-MAR-2002; 2002US-00091483.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764846.

XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Barash SC;

KW urogenital cancer; immune disorder; Addison's disease; allergy;
KW autoimmune haemolytic anaemia; autoimmune thyroiditis; diabetes mellitus;
KW Crohn's disease; multiple sclerosis; rheumatoid arthritis;
KW ulcerative colitis; acquired immunodeficiency syndrome; AIDS;
KW cardiovascular disorder; myocardial ischaemia; wound healing;
KW neurological disorder; Parkinson's disease; Alzheimer's disease;
KW cerebral anoxia; epilepsy; viral infection; bacterial infection;
KW fungal infection; parasitic infection; agonist; antagonist; ss.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT CDS 259..667
FT /*tag= a
FT /product= "Cancer related protein 4"
FT sig_peptide 259..319
FT /*tag= b
FT mat_peptide 320..664
FT /*tag= c
FT /label= Mature_Cancer_related_protein_4
XX
PN WO200118014-A1.
XX
XX
PD 15-MAR-2001.
XX
XX
PF 30-AUG-2000; 2000WO-US023794.
XX
XX
PR 03-SEP-1999; 99US-0152296P.
PR 06-OCT-1999; 99US-0158003P.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Roschke V;
XX
XX
DR WPI; 2001-235186/24.
DR P-PSDB; AAU00868.
XX

PT Twenty nine nucleic acid molecules encoding human cancer associated
PT proteins, useful in the prevention, treatment and diagnosis of cancer,
PT immune disorders, cardiovascular disorders and neurological diseases.
XX
PS Disclosure; Page 385; 427pp; English.
XX
CC The sequence encodes a novel Human cancer related protein. The
CC polynucleotides and polypeptides are useful for preventing, treating or
CC ameliorating a medical condition in e.g. humans, mice, rabbits, goats,
CC horses, cats, dogs, chickens or sheep. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities. The polynucleotide are useful for chromosome
CC identification. The nucleic acids, protein, antibodies, agonists and
CC antagonists are useful in the diagnosis, treatment and prevention of
CC cancer (e.g. cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital), immune disorders
CC (e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,
CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
CC sclerosis, rheumatoid arthritis and ulcerative colitis, acquired
CC immunodeficiency syndrome, AIDS), cardiovascular disorders such as
CC myocardial ischaemias, wound healing, neurological diseases (e.g.
CC Parkinson's disease, Alzheimer's disease, cerebral anoxia and epilepsy)
CC and infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Numerous examples of each type of disorder are given in the
CC specification
XX
SQ Sequence 1557 BP; 412 A; 345 C; 315 G; 480 T; 0 U; 5 Other;

Query Match 8.0%; Score 110.2; DB 5; Length 1557;
Best Local Similarity 90.1%; Pred. No. 6.5e-13;
Matches 118; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 1241 AAGAAAGATCTGGTTGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCT 1300
Db 1190 AAAAATTATTTTGTAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCT 1249

QY 1301 CTAGAGGATCCAAGCTTACGTACGGTGCATGCGACGTCATAGCTTCTTATAGTGCAC 1360
Db 1250 CTAGAGGATCCAAGCTTACGTACGGTGCATGCGACGTCATAGTGCAC 1309
QY 1361 CTAAATTCATT 1371
Db 1310 CTAAATTCAT 1320
RESULT 15
AAH33453
ID AAH33453 standard; cDNA; 701 BP.
XX
AC AAH33453;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:509.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; ss.
XX
OS Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US026524.
XX
PR 29-SEP-1999; 99US-0157137P.
PR 03-NOV-1999; 99US-0163280P.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI; 2001-235357/24.
DR P-PSDB; AAG74022.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers.
XX
PS Claim 1; Page 2588; 9803pp; English.
XX
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
CC proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene therapy
CC and vaccine production. N and P may be used in the prevention, diagnosis
CC and treatment of diseases associated with inappropriate P expression. For
CC example, N and P may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of P by expressing inactive proteins or to
CC supplement the patients own production of P. Additionally, N may be used
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent
CC sequences used in the exemplification of the present invention. N.B.
CC pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
CC to 1052, 7921 and 7922
XX
SQ Sequence 701 BP; 200 A; 164 C; 146 G; 181 T; 0 U; 10 Other;

Query Match 8.0%; Score 109.6; DB 4; Length 701;
Best Local Similarity 89.4%; Pred. No. 7.2e-13;
Matches 118; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1240 AAAGAAAGATCTGGTTGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGC 1299
Db 435 AAAGAGTTTARAAGAGTGTCTGGCAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGC 494

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 1, 2004, 16:15:10 ; Search time 1992 Seconds
(without alignments)
2974.522 Million cell updates/sec

Title: US-09-993-808B-1
Perfect score: 1372
Sequence: 1 cccacgcgtccggacgaag.....agtgtaacctaattcattc 1372

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_nam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID | Description |
|------------|-------|---------------|--------|----|-----------|--------------------|
| 1 | 301 | 21.9 | 1266 | 8 | AK103084 | AK103084 Oryza sat |
| 2 | 299 | 21.8 | 1242 | 6 | AX406686 | AX406686 Sequence |
| 3 | 219 | 16.0 | 568 | 6 | AX406682 | AX406682 Sequence |
| C 4 | 158.4 | 11.5 | 135670 | 8 | AP005002 | AP005002 Oryza sat |
| 5 | 114.8 | 8.4 | 2015 | 6 | BD136416 | BD136416 95 human |
| 6 | 108.4 | 7.9 | 2658 | 5 | DRE505990 | AJ505990 Danio rer |
| 7 | 107.2 | 7.8 | 2322 | 6 | AX670965 | AX670965 Sequence |
| 8 | 106.4 | 7.8 | 824 | 6 | BD072967 | BD072967 70 human |
| 9 | 105.6 | 7.7 | 2402 | 6 | AX670963 | AX670963 Sequence |
| 10 | 105.2 | 7.7 | 2454 | 6 | BD276004 | BD276004 62 Human |
| 11 | 104.4 | 7.6 | 1270 | 6 | BD275352 | BD275352 49 Human |
| 12 | 101 | 7.4 | 2657 | 6 | BD275987 | BD275987 62 Human |
| 13 | 97.2 | 7.1 | 1450 | 6 | BD247539 | BD247539 Methods f |
| 14 | 97 | 7.1 | 2486 | 6 | AX301825 | AX301825 Sequence |
| 15 | 95.6 | 7.0 | 1581 | 6 | AR066494 | AR066494 Sequence |
| 16 | 95.6 | 7.0 | 1581 | 6 | I60018 | I60018 Sequence 1 |
| 17 | 95 | 6.9 | 1255 | 6 | AR137827 | AR137827 Sequence |
| 18 | 95 | 6.9 | 1255 | 6 | AR170186 | AR170186 Sequence |
| 19 | 94.4 | 6.9 | 748 | 5 | AU301081 | AU301081 Cyprinus |
| 20 | 93.8 | 6.8 | 1495 | 6 | BD247537 | BD247537 Methods f |
| 21 | 93.4 | 6.8 | 493 | 6 | AX406678 | AX406678 Sequence |
| 22 | 93.2 | 6.8 | 1984 | 6 | BD230989 | BD230989 49 human |
| 23 | 93.2 | 6.8 | 1984 | 6 | AR337284 | AR337284 Sequence |
| 24 | 92.6 | 6.7 | 1464 | 6 | BD270355 | BD270355 50 human |
| 25 | 91.6 | 6.7 | 1665 | 6 | AR260531 | AR260531 Sequence |
| 26 | 91.6 | 6.7 | 1665 | 6 | AX058884 | AX058884 Sequence |
| 27 | 91 | 6.6 | 944 | 6 | AR184178 | AR184178 Sequence |
| 28 | 91 | 6.6 | 944 | 6 | BD192707 | BD192707 123 human |
| 29 | 88.8 | 6.5 | 2581 | 6 | BD130531 | BD130531 Ligand re |
| 30 | 87.6 | 6.4 | 2761 | 6 | AX100254 | AX100254 Sequence |
| 31 | 87 | 6.3 | 681 | 6 | AX406731 | AX406731 Sequence |
| 32 | 85.4 | 6.2 | 555 | 5 | OLA238010 | AJ238010 Oryzias 1 |
| C 33 | 84.6 | 6.2 | 125020 | 9 | AF429315 | AF429315 Homo sapi |
| 34 | 84 | 6.1 | 696 | 6 | AR184125 | AR184125 Sequence |
| 35 | 84 | 6.1 | 696 | 6 | BD192654 | BD192654 123 human |
| 36 | 82.4 | 6.0 | 831 | 6 | BD230968 | BD230968 49 human |
| 37 | 82.4 | 6.0 | 831 | 6 | AR337263 | AR337263 Sequence |
| C 38 | 82 | 6.0 | 1330 | 6 | AR137825 | AR137825 Sequence |
| C 39 | 82 | 6.0 | 1330 | 6 | AR170184 | AR170184 Sequence |
| 40 | 82 | 6.0 | 1653 | 6 | BD249962 | BD249962 50 human |
| 41 | 81.4 | 5.9 | 1813 | 6 | BD233720 | BD233720 31 human |
| 42 | 80.8 | 5.9 | 3325 | 6 | AX780203 | AX780203 Sequence |
| 43 | 80.6 | 5.9 | 425 | 5 | AU301063 | AU301063 Cyprinus |
| 44 | 80.6 | 5.9 | 2819 | 6 | BD211760 | BD211760 Proteins |
| 45 | 80.4 | 5.9 | 1668 | 6 | AX592621 | AX592621 Sequence |

ALIGNMENTS

RESULT 1
AK103084
LOCUS AK103084 1266 bp mRNA linear PLN 24-JUL-2003
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J033118G18, full insert sequence.
ACCESSION AK103084
VERSION AK103084.1 GI:32988293
KEYWORDS FLI_CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1
AUTHORS The Rice Full-Length cDNA Consortium, National Institute of

Agrobiological Sciences Rice Full-Length cDNA Project Team:
Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group: Otomo,Y., Murakami,K.,
Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Yoshino,M. and Hayashizaki,Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)

22752273

12869764

2 (bases 1 to 1266)

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K.,
Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,
Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K.,
Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,
Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,
Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
Yoshimura,A.

Direct Submission

Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica
rice.

URL : <http://cdna01.dna.affrc.go.jp/cdna/>

NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K.,
Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and
Yamamoto,M.

FAIS Genome Sequencing & Analysis Group: Otomo,Y., Iida,Y.,
Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,
Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S.,
Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N.,
Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H.,

Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F.,
Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K.,
Yasunishi,A. and Hayashizaki,Y.

FEATURES

source

1. 1266
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J033118G18"

ORIGIN

Query Match 21.9%; Score 301; DB 8; Length 1266;

Best Local Similarity 67.5%; Pred. No. 3e-41;

Matches 585; Conservative 0; Mismatches 220; Indels 62; Gaps 9;

| | | | |
|----|-----|--|-----|
| QY | 113 | AGGCGCGCGGCGTTGCGTCAGATGGGGAAGTACATGGCGAGTGCAGGGCGCGCAGGC | 172 |
| Db | 40 | AGGGCGCGCACGCGCGCAGATGGGCAAGTACATAGGAAGTTACAGGGGGCCACGGG | 99 |
| QY | 173 | GCGGAGGTCGCGCGCGTCAGGTTACGCAGGTCGTCGCGTCGCGACGAGGTCCAGTCC | 232 |
| Db | 100 | GAGGAGTTGGCGCGCATGAGGTCACGCAGGTGGTGGCGTCCGGACGAGGTCGAGTCC | 159 |
| QY | 233 | GCGGCGCGCACCGCGG-----GTCTCGGAAGGTCCGCCCGCAGGAGG | 274 |
| Db | 160 | GCAGCGCGCGCGCGCGCGACGACGACGAAGGTGAAGCGCGCTCGCGCGCTCCACGAG | 219 |
| QY | 275 | AAGAGGCGCGCGCGCGCGCGCTGCTGCCCGCGTGAGCGCTGGTGGGACGCGCGGAAGC | 334 |
| Db | 220 | AGAGGAAGCGCGTCTGCCACGCGCGTCTGCGGACTACTCGCCGTGACGCGCGGAGC | 279 |
| QY | 335 | TGCTACATCCACCTGCTAGCGCGCATGCTGTTCATGGCACCGCCTCAGCCGCGCGTCC | 394 |
| Db | 280 | TGCTACCTCCAGCTGAGGAGCGCATGCTGTTCATGGCCCGCC---GAGGCCCGCGCG | 336 |
| QY | 395 | GTTGACTCGGTTCCGACCCCGGTGGAGGCTGCTGATGGCGCTGCAGGACAGCGGCGCG | 454 |
| Db | 337 | GCGCGAGGGCTCCGGTTCTAGCGGAGGCGCGGTTCCGGGAACGAGCGCGCGCAT | 396 |
| QY | 455 | GCGCTCGCGCGCGGCTCTCGGTTGCTCCAGCACGCGCGTCTGTTGAACCTTGGGCTTG | 514 |
| Db | 397 | GCG---GCGGCTGGCTCTCGGTTGCTCCAGCACGCGCGTCTC-----CGTGGACGCG | 447 |
| QY | 515 | GGGGGTCAGCGCGGAGCCACACTGCCGCTCCTACGACGTGCAGAGGCTGGCGGGAT | 574 |
| Db | 448 | GCGGCTCAGGACAGGAGCCTCGCTGCCGCTC---CGAGCTCGCGGAGGCGAGCGGAG | 504 |
| QY | 575 | CACGCTCTGGTGGATGTCTCGGCGGCGAGCAACTCCGGAGAGCGGCCCCAGACCGGAGG | 634 |
| Db | 505 | CATGTCGCGGAGG-----GCTCGGAGCGACTCGGCGAGCGGCGCTGACCGCGGAGG | 558 |
| QY | 635 | CGAGACGACGCCATCGAGCGCGCGCGACGCGGAGCTCAGCGATCTGGAGTCGGATCTG | 694 |
| Db | 559 | AGAGAAACAACCTCCATCAAGCTTTCTCCCGGAGAGGTGAGCGATCTGGAGTCGGATCTG | 618 |
| QY | 695 | GCGGGGCACAAGA-----CTGGCCCGTCTCGTACCGGCGGCAACGCGCG | 736 |
| Db | 619 | GCTGGAGGACAGAAGCGCAGCGCTCCACTACCTTCTCGGCAACAGCCTCAGCACAGCAA | 678 |
| QY | 737 | GCTGGGAGCTGATCGTCCCGCAGCACACGAGATCCAGAGTTCTTCGCCCGCGCGGAG | 796 |
| Db | 679 | GCCACGCGGCGGAAGATTCCCGCGCGCGCGGAGATCGAGGCGTCTTCGCGCGCGCGGAG | 738 |
| QY | 797 | GCGGCCAGGCCAAGCGTTTGCTTCCAAGTACAACCTTCGACTTCGTCGCGCGCGTGC | 856 |
| Db | 739 | GAGGCTGAGGCCAAGCGTTTCGCGCCAAGTACAACCTTCGAGCTCGTTCGCGCGCGTGC | 798 |
| QY | 857 | CTCGACCGCGCGCGGTTTCGAGTGGCGCGCGGTGGTGGTTCAGCATCTGAAGC--GAGCGTG | 914 |
| Db | 799 | CTCGACGC---CGGTCGGTTCGAGTGACTCCGGTGGTTCAGCAGCGGAGGCTGAAGCGAG | 855 |

QY 915 CGTCCGCTGAAGGTGAAGCTAGAAAG 941
| | | | |
Db 856 CGTGACAGATTAAGCGGAAGCTAGAAAG 882
| | | | |
RESULT 2
AX406686
LOCUS 1242 bp
DEFINITION Sequence 9 from Patent WO0228893.
ACCESSION AX406686
VERSION AX406686.1 GI:21439634
KEYWORDS
SOURCE
ORGANISM
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1
AUTHORS Droual, A.-M., Frankard, V.M.-N., Inze, D., Mironov, V., Peres, B.A.D.
and Hatzfeld, Y.
TITLE Plant cyclin-dependent kinase inhibitors
JOURNAL Patent: WO 0228893-A 9 11-APR-2002;
CROPDESIGN N V (BE)
FEATURES
source
1. .1242
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"
ORIGIN
Query Match 21.8%; Score 299; DB 6; Length 1242;
Best Local Similarity 68.0%; Pred. No. 6.5e-41;
Matches 577; Conservative 0; Mismatches 210; Indels 62; Gaps 9;
QY 131 CAGATGGGGAAGTACATGCCCAAGTGCAGGGGCGCGGAGGTCGCGCGCTC 190
| | | | |
Db 15 CAGATGGGCAAGTACATGAGGAAGTTTCAGGGGGCCACGGGGAGGAGTTGGCCGCATG 74
| | | | |
QY 191 GAGGTTACGACGAGTCTCGGCGCTCCGACGAGGTCAGGTCGCGGCGGACCGCGC 248
| | | | |
Db 75 GAGGTCACGACGAGTGTGGCTCGGACGAGGTCGAGGTCGCGGCGGCGGCGCG 134
| | | | |
QY 249 -----GTGTCGCGAAGGTGCGCCCGGAGGAGGAGGCGCGCGCGGG 292
| | | | |
Db 135 ACGACGACGAAGGTGACGGCGGCTCGCGCGCTCCACGAGGAGGAGGCGCTGCTG 194
| | | | |
QY 293 GAGCCTGTGCGCCCGTGAGCGCTGTTGGGACGCGGGAAGCTGCTACATCCACTGCGT 352
| | | | |
Db 195 CCGACGGCGGTGCTGGGACTACTCGCGGTGACGGCGGAGCTGTACCTCCAGCTGAGG 254
| | | | |
QY 353 AGCCGCATGCTGTTCATGGCACCGCCTCAGCCGACGCGCTCGGTTGACTCGGTTCCGACC 412
| | | | |
Db 255 AGCCGCATGCTGTTCATGGCCCCGCC---GAGGCCGCGCGCGCGGAGGCTCCGTT 311
| | | | |
QY 413 CCGGTGAGGCTGTATGGCGCTGCAGGACAGAGGGCGCGCGCTCGCGCGGGCTC 472
| | | | |
Db 312 GTAGCGAGGCGCGGGTTCCGGGAACGAGCGGCGCGCATGCG---GCGGCTGGCCTC 368
| | | | |
QY 473 TCGGTTGCTCCAGCACGGCGCTGCTCGGTAACTTGGGCTTGGGGGTGTCAGCGGGGAGC 532
| | | | |
Db 369 TCGGTTGCTCCAGCACGGCGCTGCTC-----CGTGACGCGGCGGCTCAGACAGGAGC 422
| | | | |
QY 533 CACACCTGCGCTCCTACGACGCTGCAGAGGCTGCGGGGATCACGTCCTGTGATGTC 592
| | | | |
Db 423 CTCGCTGCCGCTC---CGACGTCGCGGAGGCGGAGGAGCATGTCCCGAGG----- 474
| | | | |
QY 593 TCGGCGGAGCAACTCCGGGAGCGGCCAGACCGGAGGCGGAGAGACGACGCCATCG 652
| | | | |
Db 475 -GCTCCGAGCGACTCGCGGAGCGGCGCTGACCGGAGGAGGAGAGAACTCCATCA 533
| | | | |
QY 653 AGCCGGGCGACGCGGAGCTCAGCGATCTGGAGTCGGATCTGGCGGGGACAAAGA----- 707
| | | | |
Db 534 AGCTTTCTCCCGGAGGTGAGCGATCTGGAGTCGGATCTGGCTGGAGGACAGGCGC 593
| | | | |

QY 708 -----CTGGCCCGTCTGCTACCGGCGGCAACGCGGCTGCGGAGCTGATCGTG 754
| | | | |
Db 594 AGCGTCCACTACCTTCTGCGCAACAGCCTCAGCACAGCAAGCACGCGCGGAAGATT 653
| | | | |
QY 755 CCGCAGCACACGAGATCCAGAGTCTTCGCGCGCGCGCGGCGGCGGCGGCGGCGG 814
| | | | |
Db 654 CCGCGGCGCGGAGATCGAGCGCTTCTTCGCGCGCGCGGCGGCGGCGGCGGCGG 713
| | | | |
QY 815 TTTGCTTCCAAGTACAACCTTCGACTTCGTCCGCGCGGCTGCGGCGGCGGCGGCGG 874
| | | | |
Db 714 TTCGCGCGCAAGTACAACCTTCGACTTCGTTCGCGCGGCTGCGGCGGCGGCGGCGG 770
| | | | |
QY 875 TTCGAGTGGGCGCGGCTGTCAGCATCTGAAGC--GAGCGTGGTCCGTTGCAAGGTGAA 932
| | | | |
Db 771 TTCGAGTGGACTCCGCTGGTTCAGCAGCCGGAAGCTGAAGCGGAGCGTGCAGATTAAAGCGGAA 830
| | | | |
QY 933 GCTAGAAAG 941
| | | | |
Db 831 GCTAGAAAG 839
| | | | |
RESULT 3
AX406682
LOCUS 568 bp
DEFINITION Sequence 5 from Patent WO0228893.
ACCESSION AX406682
VERSION AX406682.1 GI:21439630
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1
AUTHORS Droual, A.-M., Frankard, V.M.-N., Inze, D., Mironov, V., Peres, B.A.D.
and Hatzfeld, Y.
TITLE Plant cyclin-dependent kinase inhibitors
JOURNAL Patent: WO 0228893-A 5 11-APR-2002;
CROPDESIGN N V (BE)
FEATURES
source
1. .568
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN
Query Match 16.0%; Score 219; DB 6; Length 568;
Best Local Similarity 84.6%; Pred. No. 2.5e-27;
Matches 335; Conservative 0; Mismatches 36; Indels 25; Gaps 7;
QY 563 GCTGGCGGGATCAGTCTCTGGTGGATGTCTCGCGCGGAGCAACTCCGGGAGCGGCCCCA 622
| | | | |
Db 2 GGTACGGGGATCAGTCC---CGGATGTCTGTCGCGGAGCAACTCGGGGAGCGTCCCG 58
| | | | |
QY 623 GACCGGAGAGGCGGAGAGACGACGCCA---TCGAGCCGGGCGCA---CGGCGAGCTCAGC 676
| | | | |
Db 59 GACCGGAGAGGAGAGAGACGACGCCATCTGTCGAGCCGGCGCACGCGCGGAGCTCAGC 118
| | | | |
QY 677 GATCTGGAGTCGAGTCTGGCGG---GGCAACAAGACTGGC-----CCGTCGCTACCGGCG 727
| | | | |
Db 119 GATCTGGAGTCGAGTCTGGTGGGCGGCGAGAGACTGGCTGCTGCTGCTGCGCGCGGACA 178
| | | | |
QY 728 GCAACGCGCGCTCGGAGCTGATCGTCCCGCCAGCACACGAGATCCAGGAGTTCTTCGCC 787
| | | | |
Db 179 ACAACATCGGCTCGGAGCTGATCGTGGCGCGGAGAGATCCAGGATCCAGGAATCTTCGCG 238
| | | | |
QY 788 GCCGCCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 847
| | | | |
Db 239 GCCGCCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 298
| | | | |
QY 848 GCGGTGCCCTCGACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 907
| | | | |

QY 1288 AAGGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGGGTGCGATCGGACGTACATAGCTCT 1347
|||||
Db 1574 AAGGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGGGTGCGATCGGACGTACATAGCTCT 1633
|||||
QY 1348 TCTATAGTGTCACCTAAATTCATT 1371
|||||
Db 1634 TCTATAGTGKACCTAAATTCATT 1657
|||||
RESULT 6
DRE505990 2658 bp mRNA linear VRT 12-AUG-2003
LOCUS
DEFINITION Danio rerio mRNA for heparan sulfate 6-O-sulfotransferase (hs6st gene).
ACCESSION AJ505990
VERSION AJ505990.1 GI:29150000
KEYWORDS heparan sulfate 6-O-sulfotransferase; hs6st gene.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE 1
AUTHORS Bink,R.J., Habuhi,H., Lele,Z., Dolk,E., Joore,J., Rauch,G.J., Geisler,R., Wilson,S.W., de Hertog,J., Kimata,K. and Zivkovic,D.
TITLE Heparan Sulfate 6-O-Sulfotransferase Is Essential for Muscle Development in Zebrafish
JOURNAL J. Biol. Chem. 278 (33), 31118-31127 (2003)
PUBMED 12782624
REFERENCE 2 (bases 1 to 2658)
AUTHORS Bink,R.J.
TITLE Direct Submission
JOURNAL Submitted (19-AUG-2002) Bink R.J., Netherlands Inst. for Dev. Biology, Hubrecht Laboratory, Uppsalalaan 8, 3584CT Utrecht, NETHERLANDS
FEATURES
source Location/Qualifiers
1..2658
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/chromosome="14"
1..2658
/gene="hs6st"
314..1720
/function="brain and fin development, myogenesis"
/codon_start=1
/product="heparan sulfate 6-O-sulfotransferase"
/protein_id="CAD44529.1"
/db_xref="GI:29150001"
/db_xref="SPTREMBL:Q800H9"
/translation="MDGKSNYSRLLIALLMILFFGGIVLQYICSTSDWQLHLASLSS RLGSRAPGDRNLNGAGADPDYSSDGALVRFVPRFNTTKDLSRAVDHFHKDDVIVFL HIQKTGGTTFGRLVNRNIQLERPCECHAGQKCTCYRPGKRDITLFSFSTGWSGCLH ADWTELTNCVPSFMSNRESQERRMTPSRNYYITILRDPVWRYLSEWRHVQGATWKA SKHMCDDRLPTLTLPSCYPGDDWSGCSLEEFMVCPYNLANNRQTRMLADLSLVGCYN LTVMSNQRWAMLLS AKRNLNRNMAFFGLTEYQKTYLFEHTFRLSFIAPFTQLNGT RAASVEVEPETORRIRLELNQWDVELYVARDLFLQRFQFARQQRERRARQRIQERRK LRAVKSWLGVTKAVFKPTKEPPMTEQSPAPAEKQADAEQTLESETEGQVEENWLE EDDGEIMLDYSNVEQWR"
ORIGIN
Query Match 7.9%; Score 108.4; DB 5; Length 2658;
Best Local Similarity 99.1%; Pred. No. 1.5e-08;
Matches 109; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGT 1321
|||||
Db 2531 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGT 2590
|||||
QY 1322 ACGCGTGCATCGGACGTACATAGCTCTTCTATAGTGTCACCTAAATTCATT 1371
|||||

Db 2591 ACGCGTGCATCGGACGTACATAGCTCTTCTATAGTGTCACCTAAATTCATT 2640
|||||
RESULT 7
AX670965
LOCUS AX670965 2322 bp DNA linear PAT 27-MAR-2003
DEFINITION Sequence 3 from Patent EP1277832.
ACCESSION AX670965
VERSION AX670965.1 GI:29329461
KEYWORDS
SOURCE Heliothis virescens (tobacco budworm)
ORGANISM Heliothis virescens
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE 1
AUTHORS Ebens,A.J., Johnston,S. and Breach,J.C.
TITLE Polynucleotides encoding insect ethanolamine kinase and uses therefor
JOURNAL Patent: EP 1277832-A 3 22-JAN-2003;
Genoptera, LLC (US)
FEATURES
source Location/Qualifiers
1..2322
/organism="Heliothis virescens"
/mol_type="unassigned DNA"
/db_xref="taxon:7102"
ORIGIN
Query Match 7.8%; Score 107.2; DB 6; Length 2322;
Best Local Similarity 89.8%; Pred. No. 2.3e-08;
Matches 115; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 1244 AAAGATCTGGTTTGTCTCAAAAAAAAAAAAAAAAAAGGGCGCGCTCTA 1303
|||||
Db 1830 AAAGATTTAAATAGTCAAAAAAAAAAAAAAAAAAGGCATGGCGCGCTCTA 1889
|||||
QY 1304 GAGGATCCAAAGCTTACGTACGGTGCATGCGACGTACATAGCTCTTCTATAGTGTCACCTA 1363
|||||
Db 1890 GAGGATCCAAAGCTTACGTACGGTGCATGCGACGTACATAGCTCTTCTATAGTGTCACCTA 1949
|||||
QY 1364 AATTCATT 1371
|||||
Db 1950 AATTCATT 1957
|||||
RESULT 8
BD072967
LOCUS BD072967 824 bp DNA linear PAT 27-AUG-2002
DEFINITION 70 human secretory proteins.
ACCESSION BD072967
VERSION BD072967.1 GI:22618570
KEYWORDS JP 2001514885-A/28.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 824)
AUTHORS Ruben,S.M., Young,P.E., Brewer,L.A., Ebner,R., Olse,H.S., Florence,K.A., Rosen,C.A., Duan,R., Moore,P.A., Shi,Y., Lafleur,D.W., Florence,C., Soppet,D.R., Endress,G.A., Feng,P. and Comatsacs,G.A.
TITLE 70 human secretory proteins
JOURNAL Patent: JP 2001514885-A 28 18-SEP-2001;
HUMAN GENOME SCIENCES INC
COMMENT OS Homo sapiens (human)
PN JP 2001514885-A/28
PD 18-SEP-2001
PF 18-AUG-1998 JP 2000509821
PR 19-AUG-1997 US 60/056555,19-AUG-1997 US 60/056556 PR
19-AUG-1997 US 60/056535,19-AUG-1997 US 60/056629 PR
19-AUG-1997 US 60/056369,19-AUG-1997 US 60/056628 PR
19-AUG-1997 US 60/056728,19-AUG-1997 US 60/056368 PR
19-AUG-1997 US 60/056726,16-JUN-1998 US 60/089510 PR

15-JUL-1998 US 60/092956
PI STEVEN M RUBEN,PAUL E YOUNG,LAURIE A BREWER,REINHARD EBNER, PI
HENRIK S OLSEN,
PI KIMBERLY A FLORENCE,CRAIG A ROSEN,ROSANNE DUAN,PAUL A MOORE,
PI YANGGU SHI,
PI DAVID W LAFLEUR,CHARLES FLORENCE,DANIEL R SOPPET,GREGORY A PI
ENDRESS,
PI PING FENG,GEORGE A COMATSACS
PC C12N15/09,A61K38/00,A61P43/00,C07K16/00,C12N1/15,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12P21/02,C12Q1/68,G01N33/53,G01N33/68,C12N15/00, PC
A61K37/02,
PC C12N5/00
CC n equals a,t,g,or c
CC n equals a,t,g,or c
CC n equals a,t,g,or c
FH Key Location/Qualifiers
FT source 1..824
FT /organism='Homo sapiens (human)'.
FEATURES
source Location/Qualifiers
1..824
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 7.8%; Score 106.4; DB 6; Length 824;
Best Local Similarity 97.3%; Pred. No. 3.2e-08;
Matches 107; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAGCTTACGT 1321
|||||
DB 583 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCTCTAAGGATCCAGCTTACGT 642
QY 1322 ACGGTGCATGCGACGTACATAGCTCTTCTATAGTGCACCTAAATTCATT 1371
|||||
DB 643 ACGGTGCATGCGACGTACATAGCTCTTCTATAGTGCACCTAAATTCATT 692
RESULT 9
AX670963 2402 bp DNA linear PAT 27-MAR-2003
LOCUS
DEFINITION Sequence 1 from Patent EP1277832.
ACCESSION AX670963
VERSION AX670963.1 GI:29329460
KEYWORDS
SOURCE Heliothis virescens (tobacco budworm)
ORGANISM Heliothis virescens
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE
AUTHORS
TITLE
JOURNAL
Patent: EP 1277832-A 1 22-JAN-2003;
Genoptera, LLC (US)
FEATURES
source Location/Qualifiers
1..2402
/organism="Heliothis virescens"
/mol_type="unassigned DNA"
/db_xref="taxon:7102"
ORIGIN
Query Match 7.7%; Score 105.6; DB 6; Length 2402;
Best Local Similarity 89.1%; Pred. No. 4.4e-08;
Matches 114; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1244 AAAGATCTGGTTTGTCTCAAAAAAAAAAAAAAAAAAGGGCGGCGCTCTA 1303
|||||
DB 1939 AAAGATTAAATAGTCAAAAAAAAAAAAAAAAAAGGCCATTGCGGCGCTCTA 1998
QY 1304 GAGGATCAAGCTTACGTACGCGTGCATGCGACGTACATAGCTTCTATAGTGTACCTA 1363

|||||
DB 1999 GAGGATCCAAAGCTTACGTACGCGTGCATGCGACGTACATAGCTTCTATAGTGTACCTA 2058
QY 1364 AATTCATT 1371
|||||
DB 2059 AATTCAT 2066
RESULT 10
BD276004 2454 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION 62 Human secreted proteins.
ACCESSION BD276004
VERSION BD276004.1 GI:33085772
KEYWORDS JP 2002543771-A/82.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 2454)
Birse,C.E., Mouret,P.A., Florence,K.A., Ruben,S.M.,
Komatsoulis,G.A., Ni,J., Ebner,R., W.D., Lafleur, Olsen,H.S.,
Shi,Y., Soppet,D.R., Rosen,C.A. and Young,P.E.
TITLE 62 Human secreted proteins
JOURNAL Patent: JP 2002543771-A 82 24-DEC-2002;
Human Genome Sciences Inc
COMMENT
OS Homo sapiens
PN JP 2002543771-A/82
PD 24-DEC-2002
PF 06-APR-2000 JP 2000611564
PR 09-APR-1999 US 60/128693,26-APR-1999 US 60/130991 PI
charles e birse,paul a mouret,kimberly a florence,steven m pi
ruben,
pi george a komatsoulis,jian ni,reinhard ebner,david w pi
lafleur,henrik s olsen,
pi yanggu shi,daniel r soppet,craig a rosen,paul e young cc
FH Key Location/Qualifiers
FT SITE (2317)
FT /note='n equals a,t,g, or c'.
FEATURES
source Location/Qualifiers
1..2454
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 7.7%; Score 105.2; DB 6; Length 2454;
Best Local Similarity 97.3%; Pred. No. 5.1e-08;
Matches 107; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1262 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAGCTTACGT 1321
|||||
DB 2094 AAAAAAAAAACACAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAGCTTACGT 2153
QY 1322 ACGGTGCATGCGACGTACATAGCTTCTTCTATAGTGCACCTAAATTCATT 1371
|||||
DB 2154 ACGGTGCATGCGACGTACATAGCTTCTTCTATAGTGCACCTAAATTCAT 2203
RESULT 11
BD275352 1270 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION 49 Human Secreted Proteins.
ACCESSION BD275352
VERSION BD275352.1 GI:33085120
KEYWORDS JP 2002539787-A/28.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 1270)
Komatsoulis,G., Rosen,C.A. and Ruben,S.M.
TITLE 49 Human Secreted Proteins

JOURNAL Patent: JP 2002539787-A 28 26-NOV-2002;
Human Genome Sciences Inc
COMMENT OS Homo sapiens
PN JP 2002539787-A/28
PD 26-NOV-2002
PF 16-MAR-2000 JP 2000606742
PR 23-MAR-1999 US 60/126054,10-DEC-1999 US 60/169916 PI
george komatsoulis,craig a rosen,steven m ruben CC
FH Key Location/Qualifiers
FT SITE (553)
FT SITE /note='n equals a,t,g, or c'
FT SITE (1070)
FT SITE /note='n equals a,t,g, or c'
FT SITE (1208)
FT SITE /note='n equals a,t,g, or c'.
FEATURES Location/Qualifiers
source 1..1270
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 7.6%; Score 104.4; DB 6; Length 1270;
Best Local Similarity 90.2%; Pred. No. 7e-08;
Matches 111; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 1249 TCTGGTTGTCTCAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGA 1308
Db 960 TCTACTTAACCTCTTAATGAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGA 1019
QY 1309 TCCAGCTTACGTACGGGTGCATGCGCGTCATAGCTCTTCTATAGTGTACCTAAATTC 1368
Db 1020 TCCAGCTTACGTACGGGTGCATGCGCGTCATAGCTCTTCTATAGTGTCTCNCCTAAATTC 1079
QY 1369 ATT 1371
Db 1080 AAT 1082
RESULT 12
BD275987
LOCUS 62 Human secreted proteins. 2657 bp DNA linear PAT 17-JUL-2003
DEFINITION
ACCESSION BD275987
VERSION BD275987.1 GI:33085755
KEYWORDS JP 2002543771-A/65.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2657)
AUTHORS Birse,C.E., Mouret,P.A., Florence,K.A., Ruben,S.M.,
Komatsoulis,G.A., Ni,J., Ebner,R., W,D., Lafleur, Olsen,H.S.,
Shi,Y., Soppet,D.R., Rosen,C.A. and Young,P.E.
TITLE 62 Human secreted proteins
JOURNAL Patent: JP 2002543771-A 65 24-DEC-2002;
COMMENT OS Homo sapiens
PN JP 2002543771-A/65
PD 24-DEC-2002
PF 06-APR-2000 JP 2000611564
PR 09-APR-1999 US 60/128693,26-APR-1999 US 60/130991 PI
charles e birse,paul a mouret,kimberly a florence,steven m PI
ruben,
PI george a komatsoulis,jian ni,reinhard ebner,david w PI
lafleur,henrik s olsen,
PI yanggu shi,daniel r soppet,craig a rosen,paul e young CC
FH Key Location/Qualifiers
FT SITE (179)
FT SITE /note='n equals a,t,g, or c'.
FEATURES Location/Qualifiers
source 1..2657
/organism="Homo sapiens"

/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 7.4%; Score 101; DB 6; Length 2657;
Best Local Similarity 98.1%; Pred. No. 2.6e-07;
Matches 101; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1269 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTG 1328
Db 2304 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTG 2363
QY 1329 CATGCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCATT 1371
Db 2364 CATGCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCATT 2406
RESULT 13
BD247539
LOCUS 1450 bp DNA linear PAT 17-JUL-2003
DEFINITION Methods for transforming plastids.
ACCESSION BD247539
VERSION BD247539.1 GI:33057309
KEYWORDS JP 2002531096-A/3.
SOURCE Brassica sp.
ORGANISM Brassica sp.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 1450)
AUTHORS Chaudhuri,S.
TITLE Methods for transforming plastids
JOURNAL Patent: JP 2002531096-A 3 24-SEP-2002;
COMMENT CALGENE LLC
OS Brassica sp. (mustard)
PN JP 2002531096-A/3
PD 24-SEP-2002
PF 24-NOV-1999 JP 2000585430
PR 25-NOV-1998 US 60/109892
PI SUMITA CHAUDHURI
PC Cl2N15/09,A01H5/00,Cl2N5/10,Cl2N15/00,Cl2N5/00 CC Methods
for transforming plastids
FH Key Location/Qualifiers
FT source 1..1450
FT /organism="Brassica sp."
FEATURES Location/Qualifiers
source 1..1450
/organism="Brassica sp."
/mol_type="genomic DNA"
/db_xref="taxon:3717"
ORIGIN
Query Match 7.1%; Score 97.2; DB 6; Length 1450;
Best Local Similarity 97.1%; Pred. No. 1.2e-06;
Matches 99; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1270 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTG 1329
Db 1341 AAACCAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTG 1400
QY 1330 ATGCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCATT 1371
Db 1401 ATGCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCATT 1442
RESULT 14
AX301825
LOCUS 2486 bp DNA linear PAT 30-NOV-2001
DEFINITION Sequence 4 from Patent WO0185768.
ACCESSION AX301825
VERSION AX301825.1 GI:17382883
KEYWORDS Homo sapiens (human)
SOURCE

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
AUTHORS Gu,W.
TITLE Novel g-protein coupled receptors and uses therefor
JOURNAL Patent: WO 0185768-A 4 15-NOV-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES Location/Qualifiers
source
1..2486
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS
1..1903
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD13077.1"
/db_xref="GI:17382884"
/db_xref="REMBL:CAD13077"
/translation="NTTHRESWYACRYRSIGIPGSTHASVERSQGLSLPAHPASLAAL
AASNTTASGKLELDTFSQSSLOALDLSNNAIRSIHPEAFSTLHSLVXLDTDNQLTT
LPLAGLGLMHLKLGKLNALSOAFSKDSFPKRLILEVPYAYQCCPYGMCASFFFKASGQ
WEAEDLHLDDESSKRLPLGLLARQAEHYDDELQLEMEDSKPHPSVQCSPPTGPF
KPCYLFESWGIRLAVWAIIVLLSVLNCGLVLLTVFAGGPAPLPVKFVVGAIAGANTL
TGISCGLLASVDALTFGQFSEYGARWETGLGRATGFLAVLGSEASVLLLTIAAVQCS
VSVSCVRAYGKSPSLGSVRAGVLGCLALAGLAALPLASVGEYGASPLCLPYAPPEGQ
PAALGFTVALVMNNSFCFLVAGAYIKLYCDLPRGDFEAVWDCAMVRHVAVWLIFADGL
LYCPVAFLSFASMLGLFPVTPEAVKSVLLVLPACLNPLLLYLLFNPHRDDLRRLR
PRAGDSGPLAYAAAGELEKSSCDSTQALVAFSDVDLILEASEAGRPGLFTYGFPSVT
LISCCQPGAPRLGSHCVPEGNHFGNPQPSMDGELLRLRAEGSTPAGGGLSGGGGFQF
SGLAFASHV"

ORIGIN

Query Match 7.1%; Score 97; DB 6; Length 2486;
Best Local Similarity 95.2%; Pred. No. 1.3e-06;
Matches 100; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1267 AAAAAAAAAAAAAAAAAAGGCGCGCGCTCTAGAGGATCCAAGCTTACGTACGCG 1326
Db 2381 AAGAGATTAAAAAAAAAAAAAAAAAGGCGCGCGCTCTAGAGGATCCAAGCTTACGTACGCG 2440
QY 1327 TGCATCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCATT 1371
Db 2441 TGCATCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCATT 2485

RESULT 15
AR066494
LOCUS AR066494 1581 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5850022.
ACCESSION AR066494
VERSION AR066494.1 GI:5996710
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1581)
AUTHORS Dehesh,K., Voelker,T. and Hawkins,D.
TITLE Production of myristate in plant cells
JOURNAL Patent: US 5850022-A 1 15-DEC-1998;
FEATURES Location/Qualifiers
source
1..1581
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 7.0%; Score 95.6; DB 6; Length 1581;
Best Local Similarity 96.1%; Pred. No. 2.2e-06;
Matches 98; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1270 AAAAAAAAAAAAAAAAAAGGCGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTGC 1329
Db 1474 AAGTTAAAAAAAAAAAAAAAAAGGCGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTGC 1533

QY 1330 ATGCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCATT 1371
Db 1534 ATGCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCATT 1575
Search completed: October 1, 2004, 23:19:49
Job time : 19996 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 1, 2004, 17:12:11 ; Search time 453 Seconds
(without alignments)
1680.778 Million cell updates/sec

Title: US-09-993-808B-1
Perfect score: 1372
Sequence: 1 cccacgcgtccgagcgaag.....agtgcacctaattcattc 1372

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|----------------------|-------------------|
| 1 | 95.6 | 7.0 | 1581 | 1 US-08-383-756-1 | Sequence 1, Appli |
| 2 | 95.6 | 7.0 | 1581 | 2 US-08-460-898-1 | Sequence 1, Appli |
| 3 | 95 | 6.9 | 1255 | 3 US-09-118-442-31 | Sequence 31, Appl |
| 4 | 95 | 6.9 | 1255 | 3 US-09-677-064-31 | Sequence 31, Appl |
| 5 | 93.2 | 6.8 | 1984 | 4 US-09-904-615-46 | Sequence 46, Appl |
| 6 | 91.6 | 6.7 | 1665 | 4 US-09-591-095-23 | Sequence 23, Appl |
| 7 | 91 | 6.6 | 944 | 4 US-09-227-357-122 | Sequence 122, App |
| 8 | 88.8 | 6.5 | 2581 | 2 US-09-013-634-1 | Sequence 1, Appli |
| 9 | 84 | 6.1 | 696 | 4 US-09-227-357-69 | Sequence 69, Appl |
| 10 | 82.4 | 6.0 | 831 | 4 US-09-904-615-25 | Sequence 25, Appl |
| 11 | 82 | 6.0 | 1330 | 3 US-09-118-442-29 | Sequence 29, Appl |
| 12 | 82 | 6.0 | 1330 | 3 US-09-677-064-29 | Sequence 29, Appl |
| 13 | 79.4 | 5.8 | 826 | 4 US-09-620-312D-584 | Sequence 584, App |
| 14 | 79.4 | 5.8 | 907 | 4 US-09-620-312D-585 | Sequence 585, App |
| 15 | 79.4 | 5.8 | 1200 | 4 US-09-620-312D-564 | Sequence 564, App |
| 16 | 79.4 | 5.8 | 2544 | 4 US-09-483-371-1 | Sequence 1, Appli |
| 17 | 79.4 | 5.8 | 2544 | 4 US-09-957-156-1 | Sequence 1, Appli |
| 18 | 79 | 5.8 | 955 | 4 US-09-620-312D-228 | Sequence 228, App |
| 19 | 78.8 | 5.7 | 2638 | 3 US-09-042-785A-22 | Sequence 22, Appl |
| 20 | 78.4 | 5.7 | 501 | 4 US-09-439-313-451 | Sequence 451, App |
| 21 | 78.4 | 5.7 | 501 | 4 US-09-352-616A-451 | Sequence 451, App |
| 22 | 78.4 | 5.7 | 501 | 4 US-09-636-215-451 | Sequence 451, App |
| 23 | 78.4 | 5.7 | 501 | 4 US-09-685-166A-451 | Sequence 451, App |
| 24 | 78.4 | 5.7 | 706 | 4 US-09-439-313-449 | Sequence 449, App |
| 25 | 78.4 | 5.7 | 706 | 4 US-09-352-616A-449 | Sequence 449, App |
| 26 | 78.4 | 5.7 | 706 | 4 US-09-636-215-449 | Sequence 449, App |
| 27 | 78.4 | 5.7 | 706 | 4 US-09-685-166A-449 | Sequence 449, App |

| | | | | | |
|----|------|-----|------|-----------------------|-------------------|
| 28 | 78.4 | 5.7 | 1663 | 4 US-09-620-312D-6 | Sequence 6, Appli |
| 29 | 78.4 | 5.7 | 1790 | 4 US-09-118-637A-3 | Sequence 3, Appli |
| 30 | 78 | 5.7 | 1243 | 4 US-09-620-312D-1092 | Sequence 1092, Ap |
| 31 | 77.8 | 5.7 | 1006 | 3 US-08-911-423-3 | Sequence 3, Appli |
| 32 | 77.8 | 5.7 | 3327 | 3 US-08-689-421-26 | Sequence 26, Appl |
| 33 | 77.8 | 5.7 | 3327 | 3 US-09-389-528-26 | Sequence 26, Appl |
| 34 | 77.8 | 5.7 | 3327 | 3 US-09-181-827A-26 | Sequence 26, Appl |
| 35 | 77.4 | 5.6 | 2311 | 4 US-09-489-847-123 | Sequence 123, App |
| 36 | 77 | 5.6 | 1689 | 4 US-09-053-374A-4 | Sequence 4, Appli |
| 37 | 76.6 | 5.6 | 657 | 4 US-09-620-312D-543 | Sequence 543, App |
| 38 | 76.4 | 5.6 | 1067 | 4 US-09-620-312D-853 | Sequence 853, App |
| 39 | 76 | 5.5 | 2085 | 4 US-09-620-312D-1002 | Sequence 1002, Ap |
| 40 | 76 | 5.5 | 2664 | 4 US-09-149-476-255 | Sequence 255, App |
| 41 | 76 | 5.5 | 3396 | 4 US-09-668-680-6 | Sequence 6, Appli |
| 42 | 76 | 5.5 | 3423 | 4 US-09-668-680-7 | Sequence 7, Appli |
| 43 | 75 | 5.5 | 2202 | 4 US-09-396-149-3 | Sequence 3, Appli |
| 44 | 74.8 | 5.5 | 322 | 4 US-08-956-171E-1520 | Sequence 1520, Ap |
| 45 | 74.2 | 5.4 | 2727 | 4 US-09-620-312D-218 | Sequence 218, App |

ALIGNMENTS

RESULT 1
US-08-383-756-1
; Sequence 1, Application US/08383756
; Patent No. 5654495
; GENERAL INFORMATION:
; APPLICANT: Dehesh, Katayoon
; APPLICANT: Voelker, Toni Alois
; APPLICANT: Hawkins, Deborah
; APPLICANT: Davies, Huw Maelor
; TITLE OF INVENTION: Production of Myristate in Plant Cells
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 5.1(a)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/383,756
; FILING DATE: 02-FEB-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/261,695
; FILING DATE: 16-JUN-94
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10814
; FILING DATE: 29-OCT-93
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 07/968,971
; FILING DATE: 30-OCT-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth Lassen
; REGISTRATION NUMBER: 31,845
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510
; INFORMATION FOR SEQ ID NO: 1:


```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1581 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA to mRNA
US-08-383-756-1

```

```

Query Match          7.0%; Score 95.6; DB 1; Length 1581;
Best Local Similarity 96.1%; Pred. No. 1.7e-11;
Matches 98; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 1270 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAAGCTTACGTACGCGTGC 1329
Db 1474 AAGTTAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAAGCTTACGTACGCGTGC 1533

```

```

QY 1330 ATGCGACGTATAGCTCTTCTATAGTGTACCTAAATTCATT 1371
Db 1534 ATGCGACGTATAGCTCTTCTATAGTGTACCTAAATTCATT 1575

```

```

RESULT 2
US-08-460-898-1
; Sequence 1, Application US/08460898
; Patent No. 5850022
; GENERAL INFORMATION:
; APPLICANT: Dehesh, Katayoon
; APPLICANT: Voelker, Toni Alois
; APPLICANT: Hawkins, Deborah
; APPLICANT: Davies, Huw Maelor
; TITLE OF INVENTION: Production of Myristate in Plant Cells
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 5.1(a)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,898
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/383,756
; FILING DATE: 02-FEB-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/261,695
; FILING DATE: 16-JUN-94
; CLASSIFICATION: 435
; APPLICATION NUMBER: PCT/US93/10814
; FILING DATE: 29-OCT-93
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 07/968,971
; FILING DATE: 30-OCT-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth Lassen
; REGISTRATION NUMBER: 31,845
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 111-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (916) 753-6313

```

```

; TELEFAX: (916) 753-1510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1581 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA to mRNA
US-08-460-898-1

```

```

Query Match          7.0%; Score 95.6; DB 2; Length 1581;
Best Local Similarity 96.1%; Pred. No. 1.7e-11;
Matches 98; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 1270 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAAGCTTACGTACGCGTGC 1329
Db 1474 AAGTTAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAAGCTTACGTACGCGTGC 1533

```

```

QY 1330 ATGCGACGTATAGCTCTTCTATAGTGTACCTAAATTCATT 1371
Db 1534 ATGCGACGTATAGCTCTTCTATAGTGTACCTAAATTCATT 1575

```

```

RESULT 3
US-09-118-442-31
; Sequence 31, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 1255
; TYPE: DNA
; ORGANISM: Zea mays
US-09-118-442-31

```

```

Query Match          6.9%; Score 95; DB 3; Length 1255;
Best Local Similarity 100.0%; Pred. No. 2.1e-11;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1275 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAAGCTTACGTACGCGTGCATGCG 1334
Db 1161 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAAGCTTACGTACGCGTGCATGCG 1220

```

```

QY 1335 ACATCATAGCTCTTCTATAGTGTACCTAAATTCATT 1369
Db 1221 ACATCATAGCTCTTCTATAGTGTACCTAAATTCATT 1255

```

```

RESULT 4
US-09-677-064-31
; Sequence 31, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.

```

; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706D
; CURRENT APPLICATION NUMBER: US/09/677,064
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/055,446
; PRIOR FILING DATE: 1997-08-11
; PRIOR APPLICATION NUMBER: 60/055,526
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 60/053,944
; PRIOR FILING DATE: 1997-07-28
; PRIOR APPLICATION NUMBER: 09/118,442
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 1255
; TYPE: DNA
; ORGANISM: Zea mays
US-09-677-064-31

Query Match 6.9%; Score 95; DB 3; Length 1255;
Best Local Similarity 100.0%; Pred. No. 2.1e-11;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1275 AAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAAGCTTACGTACGGTGCATGCG 1334
Db 1161 AAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAAGCTTACGTACGGTGCATGCG 1220
QY 1335 ACGTCATAGCTTCTTATAGTGTACCTAAATTCA 1369
Db 1221 ACGTCATAGCTTCTTATAGTGTACCTAAATTCA 1255

RESULT 5
US-09-904-615-46
; Sequence 46, Application US/09904615
; Patent No. 6566325
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: PZ032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
; LENGTH: 1984
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (106)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-904-615-46

Query Match 6.8%; Score 93.2; DB 4; Length 1984;
Best Local Similarity 85.2%; Pred. No. 5.9e-11;
Matches 104; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 1240 AAAGAAAGATCTGGTTTGTCTCAAAAAAAAAAAAAAAAAAAGGGCGCGC 1299
Db 1863 AAAAAAACACATTACCTTAAGAAATACAGACAAAAAAAAAAGGGCGCGC 1922
QY 1300 TCTAGAGGATCCAAAGCTTACGTACGGTGCATGCGATCATAGCTTCTTATAGTGTCA 1359
Db 1923 TCTAGAGGATCCAAAGCTTACGTACGGTGCATGCGATCATAGCTTCTTATAGGTTCA 1982

QY 1360 CC 1361
Db 1983 CC 1984

RESULT 6
US-09-591-095-23
; Sequence 23, Application US/09591095
; Patent No. 6489461
; GENERAL INFORMATION:
; APPLICANT: Froman, B. & Dehesh, K.
; TITLE OF INVENTION: Nucleic Acid Sequences encoding Proteins
; TITLE OF INVENTION: Involved in Fatty Acid Beta-Oxidation and Methods of Use.
; FILE REFERENCE: 17137/01/US
; CURRENT APPLICATION NUMBER: US/09/591,095
; CURRENT FILING DATE: 2000-06-08
; EARLIER APPLICATION NUMBER: US 60/138,162
; EARLIER FILING DATE: 1999-06-08
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Brassica sp.
US-09-591-095-23

Query Match 6.7%; Score 91.6; DB 4; Length 1665;
Best Local Similarity 95.5%; Pred. No. 1.2e-10;
Matches 105; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 1262 AAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAAGCTTACGT 1321
Db 1545 AAAAAAAAAACCAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAAGCTTACGT 1604
QY 1322 ACGGTGCATGCGACGTCTATAGCTTCTTATAGTGTACCTAAATTCAAT 1371
Db 1605 ACGCG-GCATGCGACGTCTATAGCTTCTTATAGTGTACCTAAATTCAAT 1653

RESULT 7
US-09-227-357-122
; Sequence 122, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: PZ010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918

| | Query Match | 6.6%; | Score 91; | DB 4; | Length 944; |
|----|-----------------------|---|--------------------|----------------|-------------------|
| | Best Local Similarity | 68.5%; | Pred. No. 1.4e-10; | | |
| | Matches 139; | Conservative | 1; | Mismatches 61; | Indels 2; Gaps 1; |
| QY | 1149 | AATCCGCTTACTGTGTAATTCTCTCATGCATCTCTTAGATGGAGTTTAAATCGTCCTTAAT | 1208 | | |
| Dd | 724 | AAAGCAATTAGTGAAGCAGCTCTATCCAAAATGACITTTTGTCCTTTTTTAAAAACCAAT | 783 | | |
| QY | 1209 | TTATTACTGTACAGCAGCTTGSTTGGCTTGCAAAGAAGATCTGGTTTGTCTCAAAAAA | 1268 | | |

| | Query Match | 6.5%; | Score 88.8; | DB 2; | Length 2581; |
|----|-----------------------|---|--------------------|-----------|--------------|
| | Best Local Similarity | 88.9%; | Pred. No. 5.5e-10; | | |
| | Matches 96; | Conservative 0; | Mismatches 12; | Indels 0; | Gaps 0; |
| QY | 1241 | AAGAAAGATCTGGTTTGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCT | 1300 | | |
| Db | 2471 | AATAAAAAACCAAGCAGACGCAAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCT | 2530 | | |
| QY | 1301 | CTAGAGGATCCAAAGCTTACGTACGGTGCATGCGACGTCATAGCTCTT | 1348 | | |
| Dh | 2531 | CTAGAGGATCCAAAGCTTACGTACGGTGCATGCGACGTCATAGCTCTT | 2578 | | |

RESULT 9
US-09-227-357-69
; Sequence 69, Application US/09227357

Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12

NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 69
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (605)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (648)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (655)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-227-357-69

Query Match 6.1%; Score 84; DB 4; Length 696;
Best Local Similarity 92.2%; Pred. No. 4e-09;
Matches 95; Conservative 2; Mismatches 5; Indels 1; Gaps 1;
QY 1270 AAAAAAAAAAAAAAAAAAGGGGGGGGGCTCTAGAGGATCCAAAGCTTACGTACGGTGC 1329
|||||
Db 567 AAAAAAAAAAAAAAAAAAGGGGGGGGGCTCTAKAGNTCCAAAGCTTACGTACGGTGC 626
|||||
QY 1330 ATGGGACGTATAGCTCTCTCTAT-AGTGTACCTAAATTCATT 1371
|||||
Db 627 ATGGGACGTATAGCTCTCTCTNTAAGGNCACCTAAATTCATT 669
|||||

RESULT 10
US-09-904-615-25
; Sequence 25, Application US/09904615
; Patent No. 6566325
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: PZ032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 831
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (5)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (10)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (11)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (15)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (27)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-904-615-25

QY 1351 ATAGTGTACCTAAATTCATT 1371
Db 803 ATAGTGTACCTAAATTCAT 823

RESULT 14

US-09-620-312D-585
; Sequence 585, Application US/09620312D
; Patent No. 6569662

GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides

; FILE REFERENCE: 784CIP2B

; CURRENT APPLICATION NUMBER: US/09/620,312D

; CURRENT FILING DATE: 2000-07-19

; PRIOR APPLICATION NUMBER: 09/552,317

; PRIOR FILING DATE: 2000-04-25

; PRIOR APPLICATION NUMBER: 09/488,725

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 1105

; SOFTWARE: pt_FL_genes Version 1.0

; SEQ ID NO 585

; LENGTH: 907

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (225)..(695)

US-09-620-312D-585

Query Match 5.8%; Score 79.4; DB 4; Length 907;
Best Local Similarity 98.8%; Pred. No. 4.2e-08;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1291 GCGGGCGCTCTAGAGATCCAAAGCTTACGTACGGGTGCATGCGACGTATAGCTCTTCT 1350
Db 824 GCGGGCGCTCTAGAGATCCAAAGCTTACGTACGGGTGCATGCGACGTATAGCTCTTCT 883

QY 1351 ATAGTGTACCTAAATTCATT 1371
Db 884 ATAGTGTACCTAAATTCAT 904

RESULT 15

US-09-620-312D-564/c
; Sequence 564, Application US/09620312D
; Patent No. 6569662

GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.

; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides

; FILE REFERENCE: 784CIP2B

; CURRENT APPLICATION NUMBER: US/09/620,312D

; CURRENT FILING DATE: 2000-07-19

; PRIOR APPLICATION NUMBER: 09/552,317

; PRIOR FILING DATE: 2000-04-25

; PRIOR APPLICATION NUMBER: 09/488,725

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 1105

; SOFTWARE: pt_FL_genes Version 1.0

; SEQ ID NO 564

; LENGTH: 1200

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (559)..(957)

US-09-620-312D-564

Query Match 5.8%; Score 79.4; DB 4; Length 1200;
Best Local Similarity 98.8%; Pred. No. 4.5e-08;
Matches 80; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1291 GCGGGCGCTCTAGAGATCCAAAGCTTACGTACGGGTGCATGCGACGTATAGCTCTTCT 1350
Db 86 GCGGGCGCTCTAGAGATCCAAAGCTTACGTACGGGTGCATGCGACGTATAGCTCTTCT 27

QY 1351 ATAGTGTACCTAAATTCATT 1371
Db 26 ATAGTGTACCTAAATTCAT 6

Search completed: October 2, 2004, 05:36:30
Job time : 458 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 1, 2004, 16:15:10 ; Search time 17932 Seconds
(without alignments)
2284.793 Million cell updates/sec

Title: US-09-993-808B-1
Perfect score: 1372
Sequence: 1 cccacgcgtccgacgaag.....agtgtaacctaattcattc 1372

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

| | |
|-----|--------------|
| 1: | em_estba:* |
| 2: | em_esthum:* |
| 3: | em_estin:* |
| 4: | em_estmu:* |
| 5: | em_estov:* |
| 6: | em_estpl:* |
| 7: | em_estro:* |
| 8: | em_htc:* |
| 9: | gb_est1:* |
| 10: | gb_est2:* |
| 11: | gb_htc:* |
| 12: | gb_est3:* |
| 13: | gb_est4:* |
| 14: | gb_est5:* |
| 15: | em_estfun:* |
| 16: | em_estom:* |
| 17: | em_gss_hum:* |
| 18: | em_gss_inv:* |
| 19: | em_gss_pln:* |
| 20: | em_gss_vrt:* |
| 21: | em_gss_fun:* |
| 22: | em_gss_mam:* |
| 23: | em_gss_mus:* |
| 24: | em_gss_pro:* |
| 25: | em_gss_rod:* |
| 26: | em_gss_phg:* |
| 27: | em_gss_vrl:* |
| 28: | gb_gss1:* |
| 29: | gb_gss2:* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 581.8 | 42.4 | 707 | 11 | AY110319 Zea mays |
| 2 | 495.6 | 36.1 | 526 | 14 | CD444277 EL01N0438 |
| 3 | 455.2 | 33.2 | 1197 | 11 | AY108168 Zea mays |
| 4 | 447.6 | 32.6 | 869 | 28 | BZ743686 OGFAC93TC |

| | | | | | | |
|---|----|-------|------|------|----|----------|
| C | 5 | 441.8 | 32.2 | 666 | 29 | CC610974 |
| C | 6 | 432.8 | 31.5 | 838 | 28 | CC440801 |
| C | 7 | 392.8 | 28.6 | 505 | 14 | CA402478 |
| C | 8 | 384.2 | 28.0 | 659 | 13 | CA074449 |
| C | 9 | 375.6 | 27.4 | 885 | 29 | CG324875 |
| | 10 | 375.6 | 27.4 | 887 | 29 | CG223296 |
| | 11 | 375.2 | 27.3 | 705 | 13 | CA106521 |
| C | 12 | 374.4 | 27.3 | 655 | 13 | CA104638 |
| C | 13 | 361.2 | 26.3 | 790 | 28 | CC390573 |
| C | 14 | 361.2 | 26.3 | 835 | 29 | CG218361 |
| C | 15 | 356.6 | 26.0 | 825 | 29 | CG265472 |
| | 16 | 349.4 | 25.5 | 741 | 29 | CC683397 |
| | 17 | 348.8 | 25.4 | 1010 | 28 | CC390569 |
| | 18 | 347.6 | 25.3 | 661 | 14 | CA256119 |
| | 19 | 289.2 | 21.1 | 585 | 13 | CA104568 |
| | 20 | 283.6 | 20.7 | 584 | 13 | CA108352 |
| | 21 | 282 | 20.6 | 563 | 13 | CA108367 |
| | 22 | 272.4 | 19.9 | 510 | 14 | CB929335 |
| C | 23 | 272.4 | 19.9 | 642 | 13 | CA104689 |
| | 24 | 256.4 | 18.7 | 531 | 14 | CA297189 |
| C | 25 | 254.2 | 18.5 | 884 | 29 | CG223284 |
| | 26 | 252 | 18.4 | 488 | 28 | BH784474 |
| C | 27 | 248.8 | 18.1 | 625 | 29 | CG136785 |
| C | 28 | 242.4 | 17.7 | 795 | 29 | CG339703 |
| | 29 | 238 | 17.3 | 753 | 14 | CB646893 |
| | 30 | 236.2 | 17.2 | 787 | 28 | BZ404090 |
| C | 31 | 228 | 16.6 | 708 | 28 | BZ404098 |
| C | 32 | 219 | 16.0 | 568 | 9 | AI737717 |
| C | 33 | 216.2 | 15.8 | 550 | 12 | BG267004 |
| | 34 | 203.4 | 14.8 | 697 | 14 | CB646543 |
| | 35 | 199.2 | 14.5 | 890 | 29 | CG218373 |
| C | 36 | 198.8 | 14.5 | 225 | 14 | CD957606 |
| | 37 | 198.8 | 14.5 | 917 | 29 | CG265481 |
| C | 38 | 197.6 | 14.4 | 753 | 29 | CC729673 |
| | 39 | 197.6 | 14.4 | 791 | 29 | CG324885 |
| | 40 | 197.6 | 14.4 | 986 | 28 | CC002416 |
| C | 41 | 197.6 | 14.4 | 987 | 28 | CC002419 |
| | 42 | 188.2 | 13.7 | 657 | 28 | BZ743690 |
| C | 43 | 187.2 | 13.6 | 547 | 28 | BH777311 |
| | 44 | 157 | 11.4 | 645 | 29 | CG136784 |
| | 45 | 151.4 | 11.0 | 644 | 29 | CC729680 |

ALIGNMENTS

| | | | | | | |
|------------|--|-------------|----------------|------|--------|-----------------|
| RESULT 1 | AY110319 | AY110319 | 707 bp | mRNA | linear | HTC 17-OCT-2002 |
| LOCUS | Zea mays | CL528_-1 | mRNA sequence. | | | |
| DEFINITION | AY110319 | | | | | |
| ACCESSION | AY110319 | | | | | |
| VERSION | AY110319.1 | GI:21214635 | | | | |
| KEYWORDS | HTC. | | | | | |
| SOURCE | Zea mays | | | | | |
| ORGANISM | Zea mays | | | | | |
| REFERENCE | | | | | | |
| AUTHORS | Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V. | | | | | |
| TITLE | Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes | | | | | |
| JOURNAL | Unpublished (2002) | | | | | |
| REFERENCE | 2 (bases 1 to 707) | | | | | |
| AUTHORS | Coe,E.H. | | | | | |
| TITLE | Direct Submission | | | | | |
| JOURNAL | Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA | | | | | |
| COMMENT | If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the | | | | | |

Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES
source
1. .707
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:632312"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus Library"
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

Query Match 42.4%; Score 581.8; DB 11; Length 707;
Best Local Similarity 86.7%; Pred. No. 1.1e-58;
Matches 615; Conservative 1; Mismatches 80; Indels 13; Gaps 1;
QY 525 GCGGAGCCACACCTGCGCTCCTACGACGCTGCAGAGCTGGGGGATCAGTCCTGG 584
Db 8 GCGGAGCCACACCTGCGCTCCGACGACGCTGCAGAGCTGGGGGATCAGTCCTGG 67
QY 585 TGGATGTCTCGGCGGCGAGCAACTCCGGAGCGGCCAGACCCGAGAGCGGAGACGA 644
Db 68 TGGTGGATGTCTCGGCGGAGCAACTCCGGAGCGGCCAGACCCGAGAGGAGACGA 127
QY 645 CGCCATCGAGCCGGCGGCGAGCGGAGCTCAGCGATCTGGAGTCGGATCTGGCGGGCACA 704
Db 128 CGCCATCGAGCCGGCGGCGAGCGGAGCTCAGCGATCTGGAGTCGGATCTGGCGGGCACA 187
QY 705 AGACTGGCCCGTCCGTACCGCGGCGCAACCGCGCTCGGAGCTGATCGTCCGCGCAGCAC 764
Db 188 AGACTGGCCCGTCCGTACCGCGGCGCAACCGCGCTCGGAGCTGATCGTCCGCGCAGCAC 247
QY 765 ACAGATCCAGGAGTCTTCGCGCGCGCGAGCGGCCAGGCCCAAGCGCTTTGCTTCCA 824
Db 248 ACAGATCCAGGAGTTCNN 307
QY 825 AGTACAACTTCGACTTCGTCGCGCGCGTGCCTCCGACCGCGCGCGCGTTCGAGTGGG 884
Db 308 AGTACAACTTCGACTTCGTCGCGCGCGTGCCTCCGACCGCGCGCGCGTTCGAGTGGG 367
QY 885 CGCCGGTGTGAGCATCTGAAGCGAGCGTGCCTCGAGCGCGCGCGCGTTCGAGTGGG 944
Db 368 CGCCGGTGTGAGCATCTGAAGCGAGCGTGCCTCGAGCGCGCGCGCGTTCGAGTGGG 427
QY 945 AAAGATGCCCCCCCCCCCCCCCCCAACAAACATAACGAGAGAGAGAGAGAGAGAGAG 1004
Db 428 AAAGATG-----NNNNNAAACAAACANNNNNNNNNNNNNNNNNNNNNNNNNNNN 474
QY 1005 TAAGCAGCTTTATATAGCTTAAGCTAACCCACCACCAATTCATCTCGTCCAAATGCATGCCT 1064
Db 475 NAAGCAGCTTTATATAGCTTAAGCTAACCCACCACCAATTCATCTCGTCCAAATGCATGCCT 534
QY 1065 TGCTTTTCTCTGGAGCTAGCAGGAGCGTAGTTATTATTATTAGTACTACTTTACTTTATTCAG 1124
Db 535 TGCTTTTCTCTGGAGCTAGCAGGAGCGTAGTTATTATTATTAGTACTACTTTACTTTATTCAG 594
QY 1125 AGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTAATTTCTCTCATGCATCTCT 1184
Db 595 AGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTAATTTCTCTCATGCATCTCT 654
QY 1185 TAGATGAGTTTAAATCGTCTTAATTTATTACTGTACAGCAGCTTGSTTG 1233
Db 655 TAGATGAGTTTAAATCGTCTTAATTTATTACTGTACAGCAGCTTGCTGG 703

CD444277
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 526)
Messing, J.
Sequencing of the maize endosperm ESTs
Unpublished (2002)
Contact: Lai, Jinsheng
Dr. Joachim Messing's lab
Waksman Institute, Rutgers University
190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
Tel: 732-445-3801
Fax: 732-445-5735
Email: jlai@waksman.rutgers.edu
Seq primer: T3.
Location/Qualifiers
1. .526
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W22"
/db_xref="taxon:4577"
/tissue_type="Endosperm of 7-23DAP"
/clone_lib="Endosperm_4"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 36.1%; Score 495.6; DB 14; Length 526;
Best Local Similarity 99.2%; Pred. No. 1.3e-48;
Matches 498; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 13 GACGAAAGCGGCTGCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 72
Db 25 GACGAAAGCGGCTGCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 84
QY 73 AGAGAAAG 132
Db 85 GAGAGAAAG 144
QY 133 GATGGGAG 192
Db 145 GATGGGAG 204
QY 193 GGTACGAG 252
Db 205 GGTACGAG 264
QY 253 CGCGAAG 312
Db 265 CGCGAAG 324
QY 313 CGCTGTGGGAG 372
Db 325 CGCTGTGGGAG 384
QY 373 ACCGCTCAGCCGAG 432
Db 385 ACCGCTCAGCCGAG 444
QY 433 CGCTGAG 492
Db 445 CGCTGAG 504
QY 493 GTCGTGGTGAACCTTGGGCTTG 514

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 869)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.

Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGFAC93TM
Contact: Cathy Whitelaw

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF

Class: sheared ends.

Location/Qualifiers

1. .869

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBMA0228P17"

/clone_lib="ZM_0.7_1.5_KB"

/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb

methylation filtered genomic DNA library"

RESULT 5
CC610974/c

LOCUS

DEFINITION

CC610974

OGVFR67TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0551L14,
genomic survey sequence.

CC610974

CC610974.1 GI:31972395

GSS.

Zea mays

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 666)

Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.

Consortium for Maize Genomics
Unpublished (2002)

Other_GSSs: OGVFR67TH

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TF

Class: sheared ends.

Location/Qualifiers

1. .666

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBMA0551L14"

/clone_lib="ZM_0.7_1.5_KB"

/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb

methylation filtered genomic DNA library"

QY 429 ATGGCGTGCAGACAGCAGGCGCGGCTCGCGCGGCTCTCGCGTGTGCTCCAGCA 488
 Db 175 ATGGCGTGCAGACAGCAGGCGCGGCTCGCGCGGCTCTCGCGTGTGCTCCAGCA 116
 QY 489 CGGCGTGCCTGAACTTGGGCTTGGGGGTTCAGCGCGGAGCCACACCTGCCGCTCCT 548
 Db 115 CGGCGTGCCTGAACTTGGGCTTGGGGGTTCAGCGCGGAGCCACACCTGCCGCTCCG 56
 QY 549 ACGACGCTGCAGAGG 563
 Db 55 ACGACGCTGCAGAGG 41

RESULT 6
 CC40801/c
 LOCUS
 DEFINITION
 CC40801
 VERSION
 CC40801.1 GI:30942238
 KEYWORDS
 GSS.
 SOURCE
 Zea mays
 ORGANISM
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 838)
 Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
 Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
 Bennetzen,J.
 Maize Genomics Consortium
 Unpublished (2003)
 Other_GSSs: PUHLN53TD
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TR
 Class: sheared ends.

FEATURES
 source
 1..838
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBta478J10"
 /clone_lib="ZM_0.6_1.0 KB"
 /note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
 CoT selected genomic DNA library"

ORIGIN
 Query Match 31.5%; Score 432.8; DB 28; Length 838;
 Best Local Similarity 95.6%; Pred. No. 2.2e-41;
 Matches 480; Conservative 0; Mismatches 12; Indels 10; Gaps 3;
 QY 13 GACGCAAGCGGCTGACGAGCAGCGCGCGCGGCTGTGGCC-TGTGGGAGAGGAAA 71
 Db 512 GACGCAAGCGGCTGACGAGCAGCGCGCGCGGCTGTGGCC-TGTGGGAGAGGAAA 453
 QY 72 AAGAGAAAGAGGAAACCGGCCAAGCAAGCAAGCAGAGCGCGCGGCGGTGCGTC 131
 Db 452 AAGAGAAAGAGGAAACCGGCCAAGCAAGCAAGCAGAGCGCGCGGCGGTGCGTC 393
 QY 132 AGATGGGAAGTACATCGCAAGTGCAGGGCGCGCGAGGCGCGGAGGTGCGCGCGTCG 191
 Db 392 AGATGGGAAGTACATCGCAAGTGCAGGGCGCGCGAGGCGCGGAGGTGCGCGCGTCG 333
 QY 192 AGGTTACGAGGTGCTCGGCGTCCGGACGAGGTCCAGGTCCGCGGCGGCGCGCGGTG 251
 Db 332 AGGTTACGAGGTGCTCGGCGTCCGGACGAGGTCCAGGTCCGCGGCGGCGCGCGGTG 273

QY 252 TCGCGAAG---GTGCGCCCGGAGGAGAGGCGCGCGGCGGAGCCTGCTGCCGCGG 308
 Db 272 TCGCGAAGGTGCTGCGCCCGGAGGAGAGGCGCGCGGCGGAGCCTGCTGCCGCGG 213
 QY 309 TGAGCGCTGGTGGGACCGCGGAGGCTGCTACATCCACCTGCGTAGCCGCGCATGCTTTCA 368
 Db 212 TGGCGCTGGTGGGACCGCGGAGGCTGCTACATCCACCTGCGTAGCCGCGCATGCTTTCA 153
 QY 369 TGGACCGCCTCAGCCGCGAGCGCGCTCGGTTGACTCGGTTCCGACCCCGGTGGAGGCTGCTG 428
 Db 152 TGGACCGCCTCAGCCGCGAGCGCGG-----TCGGTTCCGACCCCGGAGGCTGCTG 99
 QY 429 ATGGCGCTGCAGGACAGCGCGCGCGCTCGCGCGGCGCTCGCGGCTGCTGCCAGCA 488
 Db 98 ATGGCGCTGCAGGACAGCGCGCGCGCTCGCGCGGCGCTCGCGGCTGCTGCCAGCA 39
 QY 489 CGGCGTGCCTGAACTTGGG 510
 Db 38 CGGCGTGCCTGAACTTGGG 17

RESULT 7
 CA402478/c
 LOCUS
 DEFINITION
 EL01N0438A02.g Endosperm_4 Zea mays cDNA, mRNA sequence.
 ACCESSION
 CA402478
 VERSION
 CA402478.1 GI:24767334
 KEYWORDS
 EST.
 SOURCE
 Zea mays
 ORGANISM
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 505)
 Lai,J., Dey,N., Kim,C.S., Becraft,P., Larkins,B., Linton,E. and
 Messing,J.
 Sequencing of the maize endosperm ESTs
 Unpublished (2002)
 Contact: Lai, Jinsheng
 Dr. Joachim Messing's lab
 Waksman Institute, Rutgers University
 190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
 Tel: 732-445-3801
 Fax: 732-445-5735
 Email: jlai@waksman.rutgers.edu
 Seq primer: T7.

FEATURES
 Location/Qualifiers
 1..505
 /organism="Zea mays"
 /mol_type="mRNA"
 /cultivar="W22"
 /db_xref="taxon:4577"
 /tissue_type="Endosperm of 7-23DAP"
 /clone_lib="Endosperm_4"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI"

ORIGIN
 Query Match 28.6%; Score 392.8; DB 14; Length 505;
 Best Local Similarity 92.2%; Pred. No. 1.2e-36;
 Matches 437; Conservative 1; Mismatches 28; Indels 8; Gaps 2;
 QY 816 TTGCTTCCAAGTACAATTGACTTCGCTCGCGCGGTGCCCTCGACGCCGCGCGGT 875
 Db 505 TTGCTTCCAAGTACAATTGACTTCGCTCGCGCGGTGCCCTCGACGCCGCGCGGT 446
 QY 876 TCGAGTGGGCGCGGTGCTGAGCATCTGAGCGAGCGTGCCTCGGTGCAAGGTGAAGCT 935
 Db 445 TCGAGTGGGCGCGGTGCTGAGCATCTGAGCGAGCGTGCCTCGGTGCAAGGTGAAGCT 386
 QY 936 AGAAGAGAGAAAGATGCCCCCCCCCCCCCCCCCAACATAACGAGAGAGAGAAAA 995
 Db 385 AGAAGAGAGAAAGATG-----CCCCCCCCCCCCCAACATAACGAGAGAGAGAAAC 333

| | | | |
|------------|---|---|------|
| QY | 996 | CCAAACAATTAAAGAGCTTTATATAGCCTAAGCTAAGCTAACCAACCAACCACTTATCTGTCCTCAA | 1055 |
| Db | 332 | CCAACAATTAAAGAGCTTTATATAGCCTAAGCTAAGCTAACCAACCAACCACTTATCTGTCCTCAA | 273 |
| QY | 1056 | TGCATGCTTCTGCTTTCTCTGGAGCTAGCAGGAGCGTAGTATTATTAGTACTTCTTA | 1115 |
| Db | 272 | TGCATGCTTCTGCTTTCTCTGGAGCTAGCAGGAGCGTAGTATTATTAGTACTTCTTA | 213 |
| QY | 1116 | CTTATTCAGAGGTTATCTTACCCCGATAGATCAATCGCTTACTGTGTAATTTCTCTCA | 1175 |
| Db | 212 | CTTATTCAGAGGTTATCTTACCCCGATAGATCAATCGCTTACTGTGTAATTTCTCTCA | 153 |
| QY | 1176 | TGCATCTCTTAGATGGAGTTAATCGTCTTAATTTACTGTACAGCAGCTTGTGCTGGC | 1235 |
| Db | 152 | TGCATCTCTTAGATGGAGTTAATCGTCTTAATTTACTGTACAGCAGCTTGC-TGGC | 94 |
| QY | 1236 | TTGCAAGAAAGATCTGTTGTCTCAAAAAAATAAAAAAATAAAAAAATAAAAAA | 1289 |
| Db | 93 | TTGCAAGAAAGATCTGTTGTCTCTAATCCGAGTCACTGTTAGCTAAAAA | 40 |
| RESULT 8 | CA074449 | 659 bp mRNA linear EST 23-SEP-2003 | |
| LOCUS | SCEZAM1081D06.g | AM1 Saccharum officinarum cDNA clone SCEZAM1081D06 | |
| DEFINITION | 5', mRNA sequence. | | |
| ACCESSION | CA074449 | | |
| VERSION | CA074449.1 | GI:34926722 | |
| KEYWORDS | EST. | | |
| SOURCE | Saccharum officinarum | | |
| ORGANISM | Saccharum officinarum | | |
| REFERENCE | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | |
| AUTHORS | Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD | | |
| TITLE | clade; Panicoideae; Andropogoneae; Saccharum. | | |
| JOURNAL | 1 (bases 1 to 659) | | |
| COMMENT | Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P. The libraries that made SUCEST Genet. Mol. Biol. 24 (1-4), 1-7 (2001) Contact: Arruda P Centro de Biologia Molecular e Engenharia Genetica Universidade Estadual de Campinas Caixa Postal 6010, 13083-970, Campinas SP, Brazil Tel: 55 19 3788 1137 Fax: 55 19 3788 1089 Email: parruda@unicamp.br Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at http://www.bcccenter.fcav.unesp.br plate: 081 row: D column: 06 Seq primer: T7 Promoter Primer. Location/Qualifiers 1..659 /organism="Saccharum officinarum" /mol_type="mRNA" /db_xref="taxon:4547" /clone="SCEZAM1081D06" /lab_host="DH10B" /clone_lib="AM1" /note="Organ: Apical meristem and tissues surrounding of mature plants; Vector: pSport1; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [apical meristem and tissues surrounding of mature plants]. cDNA was prepared from polyA+ mRNA using SuperScript plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at http://sucet.lad.ic.unicamp.br/public" | | |
| FEATURES | source | | |
| ORIGIN | Query Match | 28.0%; Score 384.2; DB 13; Length 659; | |

| | | | |
|-----------------------|---|--|---------------------------------------|
| Best Local Similarity | 83.5%; | Pred. No. 1.le-35; | |
| Matches | 573; | Conservative | 0; Mismatches 56; Indels 57; Gaps 10; |
| QY | 82 | GGAACCGGCCAAGACAAGCAAGCAGAGAGGCCAGGCC-----GCGGCGTTGCGTCAGAT | 135 |
| Db | 1 | GGAACCGGCCAAGACAAGCAAGCAGAGAGGCCAGGCC-----GCGGCGTTGCGTCAGAT | 60 |
| QY | 136 | GGGGAAGTACATGC---GCAAGTCAGGGGCGCCGAGGTCGCGCGCGTCTGA | 192 |
| Db | 61 | GGGGAAGTACATGCAGCAAGTGCAGGGGCGCCGCGGCGAGGTCGCGCGCGTCTGA | 120 |
| QY | 193 | GGTTACGAGGTCTGTCGCGCTCCGACGAGGTCAGGTCAGGTCGCGCGCGCGAGCCGCGGTG- | 251 |
| Db | 121 | GGTCACGAGGTCTGTCGCGCTCCGACGAGGTCAGGTCAGGTCGCGCGCGCGAGCCGCGGTG | 180 |
| QY | 252 | --TCGCGAAGTTCGCCCGAGGAGGAGAGGGCGC-----CGGCGGGGA | 294 |
| Db | 181 | TGTGGCGAAGTTCGCCCGAGGAGGAGGAGAGGGCGCTGGCGCGCGCGGG | 240 |
| QY | 295 | GCCTGCTGCCCGCTGAGCGCTGGTGGGACGGCGGAGAGTCTACATCCACTGCGGTAG | 354 |
| Db | 241 | GGAGCCTGCTGCCCGCTGGAGGGGACGGCGGAGAGTCTACATCCATCTGCGCAG | 300 |
| QY | 355 | CCGCATGCTGTTTCATGSCACCGCTCAGCCGCGAGCCGCTCGGTTGACTCGGTTCCGACCCC | 414 |
| Db | 301 | CCGCATGCTGTTTCATGSCACCGCTCAGCCGCGAGCCG-----TCTGTTCCGACGCC | 351 |
| QY | 415 | GGTGAGGCTGCTGATGGCGCTGCAGGACAGCAGGGCGCGCGCTCGCGCGCGGCTCTC | 474 |
| Db | 352 | GGCGGAGGCTGCTGTTGG-----AGGACAGCAGGGCGCGCGCTCGCGCGCGGCTCTC | 405 |
| QY | 475 | GGCTGCTCCAGCACGGCTCGTCCGTGAACCTTGGGCTTGGGGGTTCAGCGCGGGAGCCA | 534 |
| Db | 406 | GGCTGCTCCAGCACGGCTCGTCCGTGGACGTCGG-----GGTTCATCAGCACAGGAGCCA | 462 |
| QY | 535 | CACCTGCCGCTCCTACAGCCTGCAGAGGCTGGCGGGGATCAGTCTCTGTTGGATGTC | 594 |
| Db | 463 | CGCTGCCGCTC---CGACGCTCGGAGGTTGGCGGNGATCACGCCCGGATG-----T | 513 |
| QY | 595 | GGCGCGGAGCAACTCCGGGAGCGGCCAGACCCGAGAGCGGAGAGACGACGCCATCGAG | 654 |
| Db | 514 | CTCGCGGAGCAACTCCGGGAGCGGCCAGACCCGAGAGAGAGAGACGACGCCATCGAG | 573 |
| QY | 655 | CCGGCGGCACGGCGAGCTCAGCGATCTGGAGTCGGATCTGGCGGGGACAAAG---ACTGG | 711 |
| Db | 574 | CCGGCGGCACGGCGAGCTCAGCGATCTGGAGTCGGATCTGGCGGGGACAAAGATACTGG | 633 |
| QY | 712 | CCCGTCGCTACCGGCGGCAACCGCGG 737 | |
| Db | 634 | GGGTGCTTCCCGGCGGCAACGCTCNG 659 | |
| RESULT 9 | CG324875/c | 885 bp DNA linear GSS 26-AUG-2003 | |
| LOCUS | OGXDN04TH ZM_0.7_1.5 KB | Zea mays genomic clone ZMMBma0661B08, | |
| DEFINITION | genomic survey sequence. | | |
| ACCESSION | CG324875 | | |
| VERSION | CG324875.1 | GI:34242141 | |
| KEYWORDS | GSS. | | |
| SOURCE | Zea mays | | |
| ORGANISM | Zea mays | | |
| REFERENCE | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | |
| AUTHORS | Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD | | |
| TITLE | clade; Panicoideae; Andropogoneae; Zea. | | |
| JOURNAL | 1 (bases 1 to 885) | | |
| COMMENT | Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N. Consortium for Maize Genomics Unpublished (2002) Other_GSSs: OGXDN04TV Contact: Cathy Whitelaw | | |

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source
Location/Qualifiers
1..885
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBMA0661B08"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 27.4%; Score 375.6; DB 29; Length 885;
Best Local Similarity 95.5%; Pred. No. 9.6e-35;
Matches 422; Conservative 1; Mismatches 5; Indels 14; Gaps 3;

QY 823 CAAGTACAACTTCGACTTCGTCGGCGGCGTCCCTCGACCGCGGCGGTCGAGTG 882
|||
Db 577 CAGGTACAACTTCGACTTCGTCGGCGGCGTCCCTCGACCGCGGCGGTCGAGTG 518
|||
QY 883 GCGCGCGGTGGTCAGCATCTGAAGCGAGCGTCCGTCGGTGAAGGTGAAGCTAGAAAGA 942
|||
Db 517 GCGCGCGGTGGTCAGCATCTGAAGCGAGCGTCCGTCGGTGAAGGTGAAGCTAGAAAGA 458
|||
QY 943 GAAAAGATGCCCCCCCCCCCCCCCCCAACAACATACGAGAGAGAAACCAACA 1002
|||
Db 457 GAAAAGATG-----CCCCCAACAACAACATACGAGAGAG-AAAACCGAACA 411
|||
QY 1003 ATTAAGCAGCTTTATATAGCCTTAAGCTAACCAACCACCATTCATCTCGTCCAATGCATGC 1062
|||
Db 410 ATTAAGCAGCTTTATATAGCCTTAAGCTAACCAACCACCATTCATCTCGTCCAATGCATGC 351
|||
QY 1063 CTGTCTTTCTCTGGAGCTAGCAGGAGCGTAGTTATTTAGTACTACTTTACTTATTC 1122
|||
Db 350 CTGTCTTTCTCTGGAGCTAGCAGGAGCGTAGTTATTTAGTACTACTTTACTTATTC 291
|||
QY 1123 AGAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTAATTTCTCTCATGCATCT 1182
|||
Db 290 AGAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTAATTTCTCTCATGCATCT 231
|||
QY 1183 CTTAGATGGAGTTTAATCGTCTTAATTTATTACTGTACAGCAGCTTGSTTGGCTGCAAA 1242
|||
Db 230 CTTAGATGGAGTTTAATCGTCTTAATTTATTACTGTACAGCAGCTTGC-TGGCTTGCAAA 172
|||
QY 1243 GAAAGATCTGGTTTGTCTCAAA 1264
|||
Db 171 GAAAGATCTGGTTTGTCTCTAA 150
|||

RESULT 10
CG223296
LOCUS
DEFINITION
CG223296 887 bp DNA linear GSS 22-AUG-2003
genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 887)

REFERENCE
AUTHORS
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.

TITLE
JOURNAL
COMMENT
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGWAU17TH
Contact: Cathy Whitelaw
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
source
Location/Qualifiers
1..887
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBMA0524C09"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 27.4%; Score 375.6; DB 29; Length 887;
Best Local Similarity 95.5%; Pred. No. 9.6e-35;
Matches 422; Conservative 1; Mismatches 5; Indels 14; Gaps 3;

QY 823 CAAGTACAACTTCGACTTCGTCGGCGGCGTCCCTCGACCGCGGCGGTCGAGTG 882
|||
Db 133 CAGGTACAACTTCGACTTCGTCGGCGGCGTCCCTCGACCGCGGCGGTCGAGTG 192
|||
QY 883 GCGCGCGGTGGTCAGCATCTGAAGCGAGCGTCCGTCGGTGAAGGTGAAGCTAGAAAGA 942
|||
Db 193 GCGCGCGGTGGTCAGCATCTGAAGCGAGCGTCCGTCGGTGAAGGTGAAGCTAGAAAGA 252
|||
QY 943 GAAAAGATGCCCCCCCCCCCCCCCCCAACAACATACGAGAGAGAAACCAACA 1002
|||
Db 253 GAAAAGATG-----CCCCCAACAACAACATACGAGAGAG-AAAACCGAACA 299
|||
QY 1003 ATTAAGCAGCTTTATATAGCCTTAAGCTAACCAACCACCATTCATCTCGTCCAATGCATGC 1062
|||
Db 300 ATTAAGCAGCTTTATATAGCCTTAAGCTAACCAACCACCATTCATCTCGTCCAATGCATGC 359
|||
QY 1063 CTGTCTTTCTCTGGAGCTAGCAGGAGCGTAGTTATTTAGTACTACTTTACTTATTC 1122
|||
Db 360 CTGTCTTTCTCTGGAGCTAGCAGGAGCGTAGTTATTTAGTACTACTTTACTTATTC 419
|||
QY 1123 AGAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTAATTTCTCTCATGCATCT 1182
|||
Db 420 AGAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTAATTTCTCTCATGCATCT 479
|||
QY 1183 CTTAGATGGAGTTTAATCGTCTTAATTTATTACTGTACAGCAGCTTGSTTGGCTGCAAA 1242
|||
Db 480 CTTAGATGGAGTTTAATCGTCTTAATTTATTACTGTACAGCAGCTTGC-TGGCTTGCAAA 538
|||
QY 1243 GAAAGATCTGGTTTGTCTCAAA 1264
|||
Db 539 GAAAGATCTGGTTTGTCTCTAA 560
|||

RESULT 11

CA106521
LOCUS
DEFINITION

CA106521 705 bp mRNA linear EST 23-SEP-2003
SCQHR1013C07.g HR1 Saccharum officinarum cDNA clone SCQHR1013C07
5', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
CA106521
CA106521.1 GI:34959828
EST.
Saccharum officinarum
Saccharum officinarum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum.


```
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1. .835
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0736B20"
/clone_lib="ZM 0.7 1.5 KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 26.3%; Score 361.2; DB 29; Length 835;
Best Local Similarity 95.0%; Pred. No. 4.7e-33;
Matches 420; Conservative 1; Mismatches 4; Indels 17; Gaps 4;

QY 823 CAAGTACAACTTCGACTTCGTCGGCGGCGTGCCTCGACGCCGCGGCGGTCGAGTG 882
Db 585 CAGGTACAACTTCGACTTCGTCGGCGGCGTGCCTCGACGCCGCGGCGGTCGAGTG 526

QY 883 GCGCGCGGTGGTCAAGCATCTGAAGCGAGCGTGCCTCGGTCGAAGTGAAGCTAGAAAAGA 942
Db 525 GCGCGCGGTGGTCAAGCATCTGAAGCGAGCGTGCCTCGGTCGAAGTGAAGCTAGAAAAGA 466

QY 943 GAAAAGATGCCCCCCCCCCCCCCCCCAACAAACATAACGAGAGAGAAAACCAAAACA 1002
Db 465 GAAAAGATG-----CCCCCCCCCAACAAACATAACGAGAGAGAAAACCAAAACA 416

QY 1003 ATTAAGCAGCTTTATATAGCTTAAGCTAACCCACCACCATTCATCTCGTCCAAATGCATGC 1062
Db 415 ATTAAGCAGCTTTATATAGCTTAAGCTAACCCACCACCATTCATCTCGTCCAA--ATGC 360

QY 1063 CTTGCTTTTCTCTGGAGCTACGAGGCGTAGTTATTTAGTACTACTTTACTTATTTC 1122
Db 359 CTTGCTTTTCTCTGGAGCTACGAGGCGTAGTTATTTAGTACTACTTTACTTATTTC 300

QY 1123 AGAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTAATTTCTCTCATGCATCT 1182
Db 299 AGAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTAATTTGTCTCATGCA--T 242

QY 1183 CTTAGATGGAGTTTAATCGTCTTAATTTATTACTGTACAGCAGCTTGSTTGGCTTGCAAA 1242
Db 241 CTTAGATGGAGTTTAATCGTCTTAATTTATTACTGTACAGCAGCTTGC-TGGCTTGCAAA 183

QY 1243 GAAAAGATCTGGTTTGTCTCAA 1264
Db 182 GAAAAGATCTGGTTTGTCTCTAA 161

RESULT 15
CG265472/c
LOCUS
DEFINITION
CG265472 825 bp DNA linear GSS 25-AUG-2003
OG2CM60TH ZM_0.7_1.5 KB Zea mays genomic clone ZMMBma0764I23,
genomic survey sequence.
ACCESSION
CG265472
VERSION
CG265472.1 GI:34177613
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 825)
AUTHORS
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
TITLE
Other GSSs: OG2CM60TV
JOURNAL
COMMENT
Contact: Cathy Whitelaw
TIGR
```

```
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1. .825
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0764I23"
/clone_lib="ZM 0.7 1.5 KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

FEATURES
source
ORIGIN
```

```
Query Match 26.0%; Score 356.6; DB 29; Length 825;
Best Local Similarity 95.4%; Pred. No. 1.6e-32;
Matches 392; Conservative 1; Mismatches 5; Indels 13; Gaps 2;

QY 823 CAAGTACAACTTCGACTTCGTCGGCGGCGTGCCTCGACGCCGCGGCGGTCGAGTG 882
Db 399 CAGGTACAACTTCGACTTCGTCGGCGGCGTGCCTCGACGCCGCGGCGGTCGAGTG 340

QY 883 GCGCGCGGTGGTCAAGCATCTGAAGCGAGCGTGCCTCGGTCGAAGTGAAGCTAGAAAAGA 942
Db 339 GCGCGCGGTGGTCAAGCATCTGAAGCGAGCGTGCCTCGGTCGAAGTGAAGCTAGAAAAGA 280

QY 943 GAAAAGATGCCCCCCCCCCCCCCCCCAACAAACATAACGAGAGAGAAAACCAAAACA 1002
Db 279 GAAAAGATG-----CCCCCAACAAACATAACGAGAGAG-AAAACCGAACA 233

QY 1003 ATTAAGCAGCTTTATATAGCTTAAGCTAACCCACCACCATTCATCTCGTCCAAATGCATGC 1062
Db 232 ATTAAGCAGCTTTATATAGCTTAAGCTAACCCACCACCATTCATCTCGTCCAAATGCATGC 173

QY 1063 CTTGCTTTTCTCTGGAGCTACGAGGCGTAGTTATTTAGTACTACTTTACTTATTTC 1122
Db 172 CTTGCTTTTCTCTGGAGCTACGAGGCGTAGTTATTTAGTACTACTTTACTTATTTC 113

QY 1123 AGAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTAATTTCTCTCATGCATCT 1182
Db 112 AGAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTAATTTGTCTCATGCATCT 53

QY 1183 CTTAGATGGAGTTTAATCGTCTTAATTTATTACTGTACAGCAGCTTGSTTG 1233
Db 52 CTTAGATGGAGTTTAATCGTCTTAATTTATTACTGTACAGCAGCTTGTGCTGG 2
```

Search completed: October 2, 2004, 04:18:56
Job time : 17939 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 1, 2004, 23:19:55 ; Search time 704 Seconds
(without alignments)
9879.895 Million cell updates/sec

Title: US-09-993-808B-1
Perfect score: 1372
Sequence: 1 cccacgcgtccgacgaag.....aggtcacctaaattcattc 1372

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3340653 seqs, 2534783454 residues

Total number of hits satisfying chosen parameters: 6681306

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|--------|---------------|--------|-------|---------------------|
| 1 | 1371.6 | 100.0 | 1372 | 9 | US-09-993-308-1 |
| 2 | 1371.6 | 100.0 | 1372 | 11 | US-09-993-808B-1 |
| 3 | 990.8 | 72.2 | 1111 | 13 | US-10-425-114-34947 |
| C | 869.6 | 63.4 | 985 | 13 | US-10-425-114-18379 |
| C | 360.4 | 26.3 | 834 | 13 | US-10-425-114-18085 |
| 6 | 351 | 25.6 | 841 | 9 | US-09-993-308-5 |
| 7 | 351 | 25.6 | 841 | 11 | US-09-993-808B-5 |
| C | 301 | 21.9 | 1671 | 17 | US-10-437-963-92632 |
| 9 | 299 | 21.8 | 1242 | 16 | US-10-333-006-9 |
| 10 | 272.4 | 19.9 | 510 | 17 | US-10-767-701-24179 |
| 11 | 219 | 16.0 | 568 | 16 | US-10-333-006-5 |
| 12 | 114.8 | 8.4 | 2015 | 13 | US-10-653-595-104 |
| 13 | 114.8 | 8.4 | 2015 | 13 | US-09-397-945-104 |
| 14 | 110.8 | 8.1 | 3313 | 9 | US-09-764-846-83 |
| | | | | | Sequence 1, Appli |
| | | | | | Sequence 1, Appli |
| | | | | | Sequence 34947, A |
| | | | | | Sequence 18379, A |
| | | | | | Sequence 18085, A |
| | | | | | Sequence 5, Appli |
| | | | | | Sequence 5, Appli |
| | | | | | Sequence 92632, A |
| | | | | | Sequence 9, Appli |
| | | | | | Sequence 24179, A |
| | | | | | Sequence 5, Appli |
| | | | | | Sequence 104, App |
| | | | | | Sequence 104, App |
| | | | | | Sequence 83, Appl |

| | | | | | | |
|----|-------|-----|------|----|--------------------|-------------------|
| 15 | 110.8 | 8.1 | 3313 | 15 | US-10-091-483-83 | Sequence 83, Appl |
| 16 | 110.2 | 8.0 | 1557 | 9 | US-09-925-297-262 | Sequence 262, App |
| 17 | 110.2 | 8.0 | 1557 | 15 | US-10-023-896-42 | Sequence 42, Appl |
| 18 | 109.6 | 8.0 | 701 | 15 | US-10-106-698-519 | Sequence 519, App |
| 19 | 108.4 | 7.9 | 149 | 10 | US-09-764-891-2570 | Sequence 2570, Ap |
| 20 | 108.4 | 7.9 | 2116 | 9 | US-09-925-297-261 | Sequence 261, App |
| 21 | 108.4 | 7.9 | 2116 | 15 | US-10-023-896-31 | Sequence 31, Appl |
| 22 | 107.4 | 7.8 | 577 | 9 | US-09-925-301-628 | Sequence 628, App |
| 23 | 107.4 | 7.8 | 1837 | 9 | US-09-925-297-146 | Sequence 146, App |
| C | 107.4 | 7.8 | 2140 | 9 | US-09-764-853-250 | Sequence 250, App |
| 25 | 107.2 | 7.8 | 2322 | 15 | US-10-194-125-3 | Sequence 3, Appli |
| C | 106.8 | 7.8 | 845 | 16 | US-10-264-049-133 | Sequence 133, App |
| 27 | 106.6 | 7.8 | 858 | 15 | US-10-106-698-201 | Sequence 201, App |
| 28 | 106.4 | 7.8 | 584 | 9 | US-09-925-299-312 | Sequence 312, App |
| 29 | 106.4 | 7.8 | 584 | 10 | US-09-925-299-312 | Sequence 312, App |
| 30 | 106.4 | 7.8 | 763 | 9 | US-09-925-299-313 | Sequence 313, App |
| 31 | 106.4 | 7.8 | 763 | 10 | US-09-925-299-313 | Sequence 313, App |
| 32 | 106.4 | 7.8 | 824 | 15 | US-10-144-929-29 | Sequence 29, Appl |
| 33 | 106.4 | 7.8 | 824 | 16 | US-10-144-929-29 | Sequence 29, Appl |
| 34 | 105.6 | 7.7 | 2402 | 15 | US-10-194-125-1 | Sequence 1, Appli |
| 35 | 105.4 | 7.7 | 418 | 9 | US-09-925-299-319 | Sequence 319, App |
| 36 | 105.4 | 7.7 | 418 | 10 | US-09-925-299-319 | Sequence 319, App |
| 37 | 105.4 | 7.7 | 627 | 10 | US-09-764-891-2584 | Sequence 2584, Ap |
| 38 | 105.4 | 7.7 | 1280 | 15 | US-10-106-698-1342 | Sequence 1342, Ap |
| 39 | 105.4 | 7.7 | 1280 | 16 | US-10-264-049-75 | Sequence 75, Appl |
| 40 | 105.2 | 7.7 | 2454 | 15 | US-10-050-704-93 | Sequence 93, Appl |
| 41 | 105.2 | 7.7 | 2454 | 17 | US-10-798-512-93 | Sequence 93, Appl |
| 42 | 104.4 | 7.6 | 453 | 9 | US-09-925-299-330 | Sequence 330, App |
| 43 | 104.4 | 7.6 | 453 | 10 | US-09-925-299-330 | Sequence 330, App |
| 44 | 104.4 | 7.6 | 540 | 9 | US-09-925-299-315 | Sequence 315, App |
| 45 | 104.4 | 7.6 | 540 | 10 | US-09-925-299-315 | Sequence 315, App |

ALIGNMENTS

RESULT 1
US-09-993-308-1
; Sequence 1, Application US/09993308
; Patent No. US20020159435A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William J.
; APPLICANT: Lowe, Keith S.
; APPLICANT: Larkins, Brian A.
; APPLICANT: Dilkes, Brian R.
; APPLICANT: Sun, Yuejin
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,308
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 60/246,349
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1372
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (134)..(904)
; OTHER INFORMATION:
US-09-993-308-1

Query Match 100.0%; Score 1371.6; DB 9; Length 1372;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCCACGCGTCCGACGCAAGCGGTGTCAGGCGAGCGCGCGCGTGTGGCCTGT 60
Db 1 CCCACGCGTCCGACGCAAGCGGTGTCAGGCGAGCGCGCGCGTGTGGCCTGT 60
QY 61 GGGAGAGGAAAAGAGAGAGGAAACCGGCCCAAGACAAGCAAGCAGGAGCCAGGCCGCC 120

| | | | |
|----|------|--|------|
| Db | 61 | GGGAGAGAAAAGAGAAAGAGAAACCGGCCAAGACAAGACGAGAGGCCAGGGCCGC | 120 |
| QY | 121 | GGCGTTGCGTCAGATGGGAAGTACATGCGCAAGTGCAGGGGCGCCGAGCGCGGAGGT | 180 |
| Db | 121 | GGCGTTGCGTCAGATGGGAAGTACATGCGCAAGTGCAGGGGCGCCGAGCGCGGAGGT | 180 |
| QY | 181 | CGCCGCCGTCGAGGTTACGACAGGTCGTCGGCGTCCGGACGAGGTCAGGTCGCGGGCGGC | 240 |
| Db | 181 | CGCCGCCGTCGAGGTTACGACAGGTCGTCGGCGTCCGGACGAGGTCAGGTCGCGGGCGGC | 240 |
| QY | 241 | GACCGGCGGTGTCGCGAAAGGTCGCCCCGAGAGAGAAAGAGGGCGCCGCGGGGAGCCTGC | 300 |
| Db | 241 | GACCGGCGGTGTCGCGAAAGGTCGCCCCGAGAGAGAAAGAGGGCGCCGCGGGGAGCCTGC | 300 |
| QY | 301 | TGCCGCCGTCGAGCGCTGGTGGGACGGCGGAAGCTGCTACATCCACCTGCGTAGCCGCAT | 360 |
| Db | 301 | TGCCGCCGTCGAGCGCTGGTGGGACGGCGGAAGCTGCTACATCCACCTGCGTAGCCGCAT | 360 |
| QY | 361 | GCTGTTTCATGGCACCGCCTCAGCCGACGCGTGGTTGACTCGGTTCCGACCCCGGTGGA | 420 |
| Db | 361 | GCTGTTTCATGGCACCGCCTCAGCCGACGCGTGGTTGACTCGGTTCCGACCCCGGTGGA | 420 |
| QY | 421 | GGCTGCTGATGGCGTGCAGGACAGCAGGGCGCGCGCTCGCGGCGCGGCTCTCGCGTTG | 480 |
| Db | 421 | GGCTGCTGATGGCGTGCAGGACAGCAGGGCGCGCGCTCGCGGCGCGGCTCTCGCGTTG | 480 |
| QY | 481 | CTCCAGCACGGCGTCGTGCGTGAACCTGGGCTTGGGGGTTCAGCGCGGAGCCACACCTG | 540 |
| Db | 481 | CTCCAGCACGGCGTCGTGCGTGAACCTGGGCTTGGGGGTTCAGCGCGGAGCCACACCTG | 540 |
| QY | 541 | CCGCTCCTACGACGTCAGAGGTCGGCGGSGATCACGTCCTGGTGGATGTCCTGGCGGC | 600 |
| Db | 541 | CCGCTCCTACGACGTCAGAGGTCGGCGGSGATCACGTCCTGGTGGATGTCCTGGCGGC | 600 |
| QY | 601 | GAGCAACTCCGGGAGCGGCCAGACCGCGAGAGCGAGAGACGACGCCATCGAGCCGGGC | 660 |
| Db | 601 | GAGCAACTCCGGGAGCGGCCAGACCGCGAGAGCGAGAGACGACGCCATCGAGCCGGGC | 660 |
| QY | 661 | GCACGGCGAGCTCAGCGATCTGGAGTCGGATCTGGCGGGGCACAAAGCTGGCCGTCGCT | 720 |
| Db | 661 | GCACGGCGAGCTCAGCGATCTGGAGTCGGATCTGGCGGGGCACAAAGCTGGCCGTCGCT | 720 |
| QY | 721 | ACCGGGCGCAACGCCGGTTCGGAGCTGATCGTGCGGCCAGCACACGAGATCCAGGAGTT | 780 |
| Db | 721 | ACCGGGCGCAACGCCGGTTCGGAGCTGATCGTGCGGCCAGCACACGAGATCCAGGAGTT | 780 |
| QY | 781 | CTTCGCCGCGCGAGCGGCCAGGCCCAAGCGCTTGTCTTCCAAGTACAACCTTCGACTT | 840 |
| Db | 781 | CTTCGCCGCGCGAGCGGCCAGGCCCAAGCGCTTGTCTTCCAAGTACAACCTTCGACTT | 840 |
| QY | 841 | CGTCCGCGCGTCCCTTCGACGCCGCGCGCGGTCGAGTGGCGCGCGGTTCAGCAT | 900 |
| Db | 841 | CGTCCGCGCGTCCCTTCGACGCCGCGCGCGGTCGAGTGGCGCGCGGTTCAGCAT | 900 |
| QY | 901 | CTGAAGCGAGCGTCCGTCGAGGTGAAGCTAGAAAGAGAAAGATGCCCGCCCCC | 960 |
| Db | 901 | CTGAAGCGAGCGTCCGTCGAGGTGAAGCTAGAAAGAGAAAGATGCCCGCCCCC | 960 |
| QY | 961 | CCCCCCCCCAACAACTAACGGAGAGAGAAAAACCAACAATTAGCAGCTTATATA | 1020 |
| Db | 961 | CCCCCCCCCAACAACTAACGGAGAGAGAAAAACCAACAATTAGCAGCTTATATA | 1020 |
| QY | 1021 | GCCTAAGCTAACCAACCACTCATCTCGTCCAAATGCATGCCCTTCTCTGGAGC | 1080 |
| Db | 1021 | GCCTAAGCTAACCAACCACTCATCTCGTCCAAATGCATGCCCTTCTCTGGAGC | 1080 |
| QY | 1081 | TAGCAGGAGCGTAGTTATTATTAGTACTACTTTACTATTTCAGAGGTTATCTTGACCCC | 1140 |
| Db | 1081 | TAGCAGGAGCGTAGTTATTATTAGTACTACTTTACTATTTCAGAGGTTATCTTGACCCC | 1140 |
| QY | 1141 | GATAGATCAATCCGCTTACTGTGTAATTTCTCTCATGCATCTCTTAGATGGAGTTAATC | 1200 |
| Db | 1141 | GATAGATCAATCCGCTTACTGTGTAATTTCTCTCATGCATCTCTTAGATGGAGTTAATC | 1200 |

```

Db      1141  GATAGATCAATCCGCTTACTGTGTAAATTTCTCTCATGCATCTCTTAGATGGAGTTTAATC 1200
QY      1201  GTCTTAATTTATTAATCTGTACAGCAGCTTGSTTGGCTTGCAAAGAAAGATCTGGTTTGTCT 1260
Db      1201  GTCTTAATTTATTAATCTGTACAGCAGCTTGSTTGGCTTGCAAAGAAAGATCTGGTTTGTCT 1260
QY      1261  CAAAAAAGGCGGCGCGCGCTCTAGAGGATCCAAGCTTACG 1320
Db      1261  CAAAAAAGGCGGCGCGCGCTCTAGAGGATCCAAGCTTACG 1320
QY      1321  TACGCGTGCGATCGCAGCTCATAGCTCTTCTATAGTGTCACTAAATTCATTC 1372
Db      1321  TACGCGTGCGATCGCAGCTCATAGCTCTTCTATAGTGTCACTAAATTCATTC 1372

RESULT 2
US-09-993-808B-1
; Sequence 1, Application US/09993808B
; Publication No. US20040003433A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William
; APPLICANT: Lowe, Keith
; APPLICANT: Sun, Yuejin
; APPLICANT: Dilkes, Brian
; APPLICANT: Larkins, Brian
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides,
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,808B
; CURRENT FILING DATE: 2001-11-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1372
; TYPE: DNA
; ORGANISM: zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (134)...(902)
US-09-993-808B-1

```

| | Query Match | 100.0%; Score 1371.6; DB 11; Length 1372; | |
|----|----------------------------|--|-----------|
| | Best Local Similarity | 100.0%; Pred. No. 0; | |
| | Matches 1372; Conservative | 0; Mismatches | 0; Indels |
| | | | 0; Gaps |
| | | | 0; |
| QY | 1 | CCACGCGTCCGGACGCAAGCGGCTGCAGGCAGACGCGCCGCGCAGGCGTTGTGGCCTGT | 60 |
| Db | 1 | CCACGCGTCCGGACGCAAGCGGCTGCAGGCAGACGCGCCGCGCAGGCGTTGTGGCCTGT | 60 |
| QY | 61 | GGGAGAGGAAAAAGAGAAAGAGGAACCGGCCAAGACAAGCAAGCGAGAGGCCAGGGCCGC | 120 |
| Db | 61 | GGGAGAGGAAAAAGAGAAAGAGGAACCGGCCAAGACAAGCAAGCGAGAGGCCAGGGCCGC | 120 |
| QY | 121 | GGCGTTGCGTCAGATGGGAAGTACATGCGCAAGTGCAGGGGCGCCGCGAGGCGCGGAGGT | 180 |
| Db | 121 | GGCGTTGCGTCAGATGGGAAGTACATGCGCAAGTGCAGGGGCGCCGCGAGGCGCGGAGGT | 180 |
| QY | 181 | CGCGCCGTCGAGGTTACGCAAGTTCGTGGCGTCCGGACCGAGGTCCAGGTCGCGGGCGC | 240 |
| Db | 181 | CGCGCCGTCGAGGTTACGCAAGTTCGTGGCGTCCGGACCGAGGTCCAGGTCGCGGGCGC | 240 |
| QY | 241 | GACCGCGGTGTCGCGAAGTTCGCCCCGAGGAGGAAGAGGGCGCCGGCGGGGAGCCTGC | 300 |
| Db | 241 | GACCGCGGTGTCGCGAAGTTCGCCCCGAGGAGGAAGAGGGCGCCGGCGGGGAGCCTGC | 300 |
| QY | 301 | TGCGCGCGTGAGCGCTGTGGGGACGGCGGAAGCTGCTACATCCACCTGCGTAGCCGCAT | 360 |
| Db | 301 | TGCGCGCGTGAGCGCTGTGGGGACGGCGGAAGCTGCTACATCCACCTGCGTAGCCGCAT | 360 |
| QY | 361 | GCTGTTTCATGGCACCGCTCAGCCGACGCCGTGGTTGACTCGGTTCCGACCCCGGTGGA | 420 |
| Db | 361 | GCTGTTTCATGGCACCGCTCAGCCGACGCCGTGGTTGACTCGGTTCCGACCCCGGTGGA | 420 |

| | | | |
|----|------|---|------|
| QY | 421 | GGCTGCTGATGCGCGTGCAGGACAGCAGGGCGCGCGCTCGCGGCCGGCTCTCGCGTTG | 480 |
| Db | 421 | | 480 |
| QY | 481 | CTCCAGCACGGCGTCGTCCGTGAAC TTGGCTTG3GGGGTTCAGCGCGGAGCCACACCTG | 540 |
| Db | 481 | | 540 |
| QY | 541 | CCGCTCCCTACGACGCTGCAGAGGCTGGCGGGATCACGTCTGTTGGATGTCTCGGCGGC | 600 |
| Db | 541 | | 600 |
| QY | 601 | GAGCAACTCCGGAGCGGCCCCAGACC CGAGAGGCGAGAGACGACGCCATCGAGCCGGC | 660 |
| Db | 601 | | 660 |
| QY | 661 | GCACGGCGAGCTCAGCGATCTGGAGTCGGATCTGGCGGGGCACAAGACTGGCCCGTCGCT | 720 |
| Db | 661 | | 720 |
| QY | 721 | ACGGCGGCAACGCCGGCTGCGGAGCTGATCTGTGCCGCCAGCACACGAGATCCAGGAGTT | 780 |
| Db | 721 | | 780 |
| QY | 781 | CTTCGCCCGCGCGAGGCGGCCAGGCCAAGCGTTTGCTTCCAAGTACAACCTTCGACTT | 840 |
| Db | 781 | | 840 |
| QY | 841 | CGTCCGGCGGTGCCCTCGACCGCGCGCGGTTCGAGTGGCGCCGCTGGTTCAGCAT | 900 |
| Db | 841 | | 900 |
| QY | 901 | CTGAAGCGAGCGTCCGTCCGGTGCAAGGTGAAGCTAGAAAGAGAAAAGATGCCCCCCCCC | 960 |
| Db | 901 | | 960 |
| QY | 961 | CCCCCCCCCAACAACATAACGGAGAGAGAAAAACCAACAATTAAGCAGCTTTATATA | 1020 |
| Db | 961 | | 1020 |
| QY | 1021 | GCCTAAGCTAACCAACCACTTCATCTCGTCCAAATGCATGCCCTTGCTTTCTCTGGAGC | 1080 |
| Db | 1021 | | 1080 |
| QY | 1081 | TAGCAGGAGCGTAGTTATTATTAGTACTACTTTACTATTATTCAGAGGTTATCTTGACCCC | 1140 |
| Db | 1081 | | 1140 |
| QY | 1141 | GATAGATCAATCCGCTTACTGTGTAATTTCTCATGTCATCTCTTAGATGGAGTTTAATC | 1200 |
| Db | 1141 | | 1200 |
| QY | 1201 | GTCTTAATTTATTACTGTACAGCAGCTTGSITGCTTGCAAGAAAGATCTGGTTTGCT | 1260 |
| Db | 1201 | | 1260 |
| QY | 1261 | CAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACG | 1320 |
| Db | 1261 | | 1320 |
| QY | 1321 | TACGCGTGATGCGACGTCTAGCTCTTCTATAGTGTCACTTAATTCATT | 1372 |
| Db | 1321 | | 1372 |

RESULT 3

US-10-425-114-34947

; Sequence 34947, Application US/10425114

Publication No. US20040034888A1

GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

QY 880 GTGGGCGCGGTGTCAGCATCTGAAGCGAGCGTCCGCTCCGCTGCAAGGTGAAGCTAGAA 939
Db 769 GTGGGCGCGGTGTCAGCATCTGAAGCGAGCGTCCGCTCCGCTGCAAGGTGAAGCTAGAA 828
QY 940 AGAGAAAAGATGCCCCCCCCCCCCCCCCCAACAAACATAACGGAGAGAGAAAAACCAA 999
Db 829 AGAGAAAAGATG-----CCCCCCCCCAACAAACATAACGGAGAGAGAAAAACCAA 878
QY 1000 ACAATTAAGCAGCTTTATATAGCCTAAGCTAACCAACCAACCATTCATCTCGTCCAAATGCA 1059
Db 879 ACAATTAAGCAGCTTTATATAGCCTAAGCTAACCAACCAACCATTCATCTCGTCCAA--A 934
QY 1060 TGCCTTGCTTTCTTGAGCTAGCAGGAGCGTAGTTATTATTAGTACTACTTTACTTA 1119
Db 935 TGCCTTGCTTTCTTGAGCTAGCAGGAGCGTAGTTATTATTAGTACTACTTTACTTA 994
QY 1120 TTCAGAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTAAATTTCTCTCATGCA 1179
Db 995 TTCAGAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTAAATTTCTCTCATGCA 1054
QY 1180 TCTCTTAGATGGAGTTAATCGTCTTAATTTAATTTACTGTACAGCAGCTTGCTG 1233
Db 1055 --TCTTAGATGGAGTTAATCGTCTTAATTTAATTTACTGTACAGCAGCTTGCTG 1106

RESULT 4
US-10-425-114-18379/c
; Sequence 18379, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 18379
; LENGTH: 985
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3079-036-F9_FLI

Query Match 63.4%; Score 869.6; DB 13; Length 985;
Best Local Similarity 95.8%; Pred. No. 1.4e-226;
Matches 966; Conservative 1; Mismatches 10; Indels 31; Gaps 6;
QY 229 GTCCGCGGCGGACCGCGGTGTGCGGAAG---GTCGCCCCGAGGAGGAGGGCGCC 285
Db 985 GTCCGCGGCGGACCGCGGTGTGCGGAAGGTGCTGCGCCCCGAGGAGGAGGGCGCC 926
QY 286 GCGCGGGGAGCCTGCTGCCCGCGTGTGCGGAGCGTGGTGGGACGCGGAGCTGCTACATCCA 345
Db 925 GCGCGGGGAGCCTGCTGCCCTCCGTGGCGCTGTGGGACGCGGAGCTGCTACATCCA 866
QY 346 CCTGCGTAGCCGATGCTGTTCATGGCACCGCCTCAGCCGCGAGCGGTGACTCGGT 405
Db 865 CCTGCGTAGCCGATGCTGTTCATGGCACCGCCTCAGCCGCGAGCGCCG---TCGGT 812
QY 406 TCCGACCCCGGTGGAGCTGTGATGGCGTGCAGGACAGAGGCGCGGCGCTCGCGGC 465
Db 811 TCCGACCCCGGTGGAGCTGTGATGGCGTGCAGGACAGAGGCGCGGCGCTCGCGGC 752
QY 466 CGGGCTCTCGGTTGCTCCAGCACGGCGTGTGCGGTGAATTTGGGCTTGGGGGTGAGCG 525
Db 751 CGGGCTCTCGGTTGCTCCAGCACGGCGTGTGCGGTGCA-----CGTGGGGGTGAGCG 698

QY 526 CGGGAGCCACACCTGCGGCTCCTACGACGCTGACAGAGGCTGGCGGGGATCAGCTCTGGT 585
Db 697 CGGGAGCCACACCTGCGGCTCCGACGACGCTGACAGAGGCTGGCGGGGATCAGCTCTGGT 638
QY 586 GGATGTCTCGGCGGCGAGCAACTCCGGGAGCGGCCAGACCCGAGAGGAGAGACGAC 645
Db 637 GGATGTCTCGGCGGCGAGCAACTCCGGGAGCGGCCAGACCCGAGAGGAGAGACGAC 578
QY 646 GCCATCGAGCGCGGCGACGGGAGCTCAGGATCTGGAGTGGATCTGGCGGGGACAA 705
Db 577 GCCATCGAGCGCGGCGACGGGAGCTCAGGATCTGGAGTGGATCTGGCGGGGACAA 518
QY 706 GACTGGCCCGTCCGTACCGGGCGCAACCGCGCTGCGGAGCTGATCGTCCGCGCAGCACA 765
Db 517 GACTGGCCCGTCCGTACCGGGCGCAACCGCGCTGCGGAGCTGATCGTCCGCGCAGCACA 458
QY 766 CGAGATCCAGGAGTTCTTTCGCGCGCGCGGAGCGGCCAGGCGCTTTGCTTCCAA 825
Db 457 CGAGATCCAGGAGTTCTTTCGCGCGCGCGGAGCGGCCAGGCGCTTTGCTTCCAA 398
QY 826 GTACAACTTCGACTTCGTCCGCGCGCTGCGCCCTCGACCGCGCGCGCTTCGAGTGGC 885
Db 397 GTACAACTTCGACTTCGTCCGCGCGCTGCGCCCTCGACCGCGCGCGCTTCGAGTGGC 338
QY 886 GCCGTGTGTAGCATCTGAAGCGAGCGGTGCGTCCGCTGCAAGGTGAAGTAGAAAGAGAA 945
Db 337 GCCGTGTGTAGCATCTGAAGCGAGCGGTGCGTCCGCTGCAAGGTGAAGTAGAAAGAGAA 278
QY 946 AAGATGCCCGCCCCCCCCCAACAAACATAACGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1005
Db 277 AAGATG-----CCCCCCCCCAACAAACATAACGGAGAGAGAGAGAGAGAGAGAGAGAG 228
QY 1006 AAGCAGCTTTATATAGCCTAAGCTAACCAACCAACCATTCATCTCGTCCAAATGCCTT 1065
Db 227 AAGCAGCTTTATATAGCCTAAGCTAACCAACCAACCATTCATCTCGTCCAA--ATGCCTT 172
QY 1066 GCTTTTCTCTGAGCTAGCAGGAGCGTAGTTATTATTAGTACTACTTACTTATTTCAGA 1125
Db 171 GCTTTTCTCTGAGCTAGCAGGAGCGTAGTTATTATTAGTACTACTTACTTATTTCAGA 112
QY 1126 GGTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTAAATTTCTCTCATGCTCTT 1185
Db 111 GGTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTAAATTTCTCTCATGCTCTT 54
QY 1186 AGATGGAGTTAATCGTCTTAATTTTACTGTACAGCAGCTTGTTG 1233
Db 53 AGATGGAGTTAATCGTCTTAATTTTACTGTACAGCAGCTTGCTG 6

RESULT 5
US-10-425-114-18085/c
; Sequence 18085, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 18085
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3078-022-G5_FLI

US-10-425-114-18085

```
Query Match      26.3%; Score 360.4; DB 13; Length 834;
Best Local Similarity 74.5%; Pred. No. 1.1e-87;
Matches 611; Conservative 0; Mismatches 146; Indels 63; Gaps 10;

QY 22 GGCTGCAGGCAGCGCCGCGCAGGCGTTGTGGCTGTGGAGAGGAAAGAGAAAGA 81
Db 814 GCCTGCAGGCAGCGCCGCGCAGGCGTTGGAGTGGCCAGTGGAGTGGAGTGAAGA 755

QY 82 GGAACCGGCCAAGCAAGCAAGCAGAGGCCAGG-----GCCCGCGCGTTGCTCAGAT 135
Db 754 GGAACCGGCCAAGCAAGCAAGCAGAGGAGGAGTGTGCGCGCGGTTCCGTAAGAT 695

QY 136 GGGGAAGTACATGCGCAAGTGCAGGGGGCGCGCAGGCGCGAGGTGCGCGCGCTCGAGGT 195
Db 694 GGGGAAGTACATGCGCAAGCAGCGGGGGCGCGCAGGCGCGAGGTGCGCGCGCTCGAGGT 635

QY 196 TACGCAGGTCTGCGCGTCCGACAGAGTCCAGTCCGCGCGCGCGAC---CGGCGGTGT 252
Db 634 CTCGCAGGTCTGCGCGTCCGACAGAGTCCAGTCCGCGCGCGCGACCGCGCGGTGT 575

QY 253 CGCGAAGGTGCGCCCGAGGAGGAAGA-----GGGCGCC 285
Db 574 CGCGAAGGTGCGTCCGCGAGGAGGAAGAAGGCGTGTGCTCCGCGCGCGAAGCAGACGCG 515

QY 286 GGGGGGAGCCTGTGCTCCGCGGTGAGCGTGTGGGAGCGCGGAA---GCTGTACAT 342
Db 514 GTCGGGGAGCCTGTGCTCCGCGTGGCGCTGTGTGGGAGCGCGGAGCTGTGCTACAT 455

QY 343 CCACCTGCGTAGCCTGCTGTTTCATGGCACCGCTCAGCCGCGAGCGTCCGTTGACTC 402
Db 454 CCACCTGCGGAGCGCATGCTGTTTCATGGCAGCATCTCAGCAGCAACCGTCCG-----C 401

QY 403 GGTTCGACCCCGTGGAGGTCTGTATGGCGCTGCAGGACAGCAGGCGCGCGCTCGC 462
Db 400 GGCTCTACGCGCGTGGAGGTCTGTGTGCGGACAGCAGGCGCGGTGTGGCGCTCGC 341

QY 463 GGCGGGCTCTCGGTTGCTCCAGCACGGCGTCTGCGGTGAACCT---GGGTTGGGG 519
Db 340 GGCTGGCTCTCGGTTGCTCCAGCACGGCGTCTGCGGTGAGCTGCGGGGCGCAGCAGCC 281

QY 520 TCAGCGGGAGCCACCTGCGCTCCTACGACGCTGCAGAGGTGCGCGGGATCACGT 579
Db 280 CGCGAGCGGGAGCCACGCTGCGCTCCGACGCTGCGCCTGCGGAGTTGACGGGATCA 221

QY 580 CCTGGTGGATGTCTCGCGCGGAGCAACTCCGGGAGCGGCCAGACCGCGAGCGCGAGA 639
Db 220 CGTCCCGGATGTCTACCGCGAGCAACTCGGGGAGCGTCCCGGACCGCGAGGAGAGA 161

QY 640 GACGACGCCA---TCGACCGGGCGCA---CGGCGAGCTCAGCGATCTGAGTCCGATCT 693
Db 160 GACGACGCCATCTGTCAGCGCGGGCGCACGGCGGCGAGCTCAGCGATCTGAGTCCGATCT 101

QY 694 GCGG---GGCAAGACTGGC-----CCGTCTGTACCGCGCGCAACCGCGCTGCGGA 744
Db 100 GGTGGCGGCGAGAACTGGTGTCTGTCTGTCGCGCGCGCAACAACACTGCGGTGCGGA 41

QY 745 GCTGATCGTCCCGCAGCACAGATCCAGGATCTTCTC 784
Db 40 GCTGATCGTCCCGCAGCACAGGAGATCCAGGAATCTTCTC 1
```

RESULT 6

```
US-09-993-308-5
; Sequence 5, Application US/09993308
; Patent No. US20020159435A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William J.
; APPLICANT: Lowe, Keith S.
; APPLICANT: Larkins, Brian A.
; APPLICANT: Dilkes, Brian R.
; APPLICANT: Sun, Yuejin
```

```
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,308
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 60/246,349
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 841
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (159)..(839)
; OTHER INFORMATION: The 'r' at location 491 stands for g or a.
US-09-993-308-5
```

```
Query Match      25.6%; Score 351; DB 9; Length 841;
Best Local Similarity 73.9%; Pred. No. 4.1e-85;
Matches 610; Conservative 1; Mismatches 136; Indels 78; Gaps 10;

QY 22 GGCTGCAGGCAGCGCCGCGCAGGCGTTGTGGCTGTGGAGAGGAAAGAGAAAGA 81
Db 41 GCCTGCAGGCAGCGCCGCGCAGGCGTGGAGTGGCGAGTGGAGTGGAGTGAAGA 100

QY 82 GGAACCGGCCAAGCAAGCAAGCAGAGGCCAGG-----GCCCGCGCGTTGCTCAGAT 135
Db 101 GGAACCGGCCAAGCAAGCAAGCAGAGGAGGAGTGTGCGCGCGCGTCCGTAAGAT 160

QY 136 GGGGAAGTACATGCGCAAGTGCAGGGGGCGCGCAGGCGCGAGGTGCGCGCGCTCGAGGT 195
Db 161 GGGGAAGTACATGCGCAAGCAGCAGGGGGCGCGCGGAGGGGTGGCCGAGTCCGAGGT 220

QY 196 TACGCAGGTCTGCGCGTCCGACAGAGTCCAGTCCGCGCGCGCGAC---CGGCGGTGT 252
Db 221 CTCGCAGGTCTGCGCGTCCGACAGAGTCCAGTCCGCGCGCGCGACCGCGCGGTGT 280

QY 253 CGCGAAGGTGCGCCCGAGGAGGAAGA-----GGGCGCC 285
Db 281 CGCGAAGGTGCTCCGCGAGGAGGAAGAAGGCGTGTGCTCCCGCGCGAAGTACGAC 340

QY 286 GGCGGGGAGCCTGTGCTCCGCGCTGAGCGCTGTGTGGGAGCGCGGAA---GCTGTACAT 342
Db 341 GTCGGGGAGCCTGTGCTCCGCTGGCGCTGTGTGGGAGCGCGGAGCTGTGCTACAT 400

QY 343 CCACCTGCGTAGCCTGCTGTTTCATGGCACCGCTCAGCCGCGAGCGCTCGTTGACTC 402
Db 401 CCACCTGCGGAGCGCATGCTGTTTCATGGCAGCATCTCAGCAGCAACCGTCCG-----C 454

QY 403 GGTTCGACCCCGTGGAGGTCTGTATGGCGCTGCAGGACAGCAGGCGCGCGCTCGC 462
Db 455 GGCTCTGACGCGGTGGAGGTGTGTGTGCGGACACARCAAGGCGGGGTGTGGCGCTCGC 514

QY 463 GGCGGGGCTCTCGGTTGCTCCAGCACGGCGTGTGCGTGAACCTTGGGCTTGGGGGTCA 522
Db 515 GGCTGGCTCTCGGTTGTCTCCAGCACGGCGTGTGCGTGGTGGAGCT----- 559

QY 523 GCGCGGAGCCACACTGCGCTCCTACGACGCTGCAGAGGTGCGCGGAGTCAAGTCTCCT 582
Db 560 ---CGGGGGCCACGCTGCGCTCCGACGCTCCGCTGCGGAGTTGACGGGATCACGT 616

QY 583 GGTGGATGTCTCGCGCGGAGCAACTCCGGGAGGGGCCAGACCGCGAGAGCGGAGAGAC 642
Db 617 CCGGATGTCTACCGCGAGCAACTCGGGGAGCGTCCCGGACCGCGAGAGAGAGAC 676

QY 643 GACGCCA---TCGAGCCGGCGCA---CGGCGAGCTCAGCGATCTGAGTCCGATCTGGC 696
Db 677 GACGCCATCTGTCAGCCGGCGCACGCGGCGAGCTCAGCGATCTGAGTCCGATCTGGT 736

QY 697 GG---GGCACAAGACTGGC-----CCGTCTGTACCGCGCGCAACCGCGCTCGGAGCT 747
Db 737 GGGGCGGCAGAACTGGTGTCTGTCTGTCGCGCGCGCAACAACATCGCTCGGAGCT 796
```


Db 930 GCCGCGAGGCTCCGGTTAGCGGAGCGCGGGTCCGGGAACGAGCGCGCGGCAT 871
QY 455 GCGCTCGCGCGCGGCTCTCGGTTGCTCCAGACGCGGCTCGTGGAATTTGGGCTTG 514
Db 870 GCG---GCGGCTGGCCTCTCGGTTGCTCCAGCACGCGGCTCGTC-----CGTGAACGCG 820
QY 515 GGGGTTAGCGCGGAGCCACACCTGCCGCTCCTACGACGCTGCAGAGGCTGGCGGGAT 574
Db 819 GCGGCTCAGGACAGGAGCCTCGGTCGCCGCTC---CGACGTCGCGGAGGCGAGCGAG 763
QY 575 CACGTCCTGGTGGATGCTCGGCGGCGAGCAACTCCGGGAGCGGCCAGACCGGAGAGG 634
Db 762 CATGTCCCGGAGG-----GCTCCGCGAGCGACTCGCGGAGCGCGCTGACCGGAGAGG 709
QY 635 CGAGAGACGACGCCATCGAGCCGGCGCACGCGGAGCTCAGCGATCTGGAGTCGGATCTG 694
Db 708 AGAGAAACAACCTCAATCAAGCTTTCTCCCGGAGAGGTGAGCGATCTGGAGTCGGATCTG 649
QY 695 GCGGGGCCACAAGA-----CTGGCCCGTCGCTACCGCGCGCAACGCGG 736
Db 648 GCTGGAGGACAGAGCGCAGCGCTCCACTACCTTCTCGGGAACAGCCTCAGCACAGCAA 589
QY 737 GCTCGGAGCTGATCGTCCGCGCCAGCACACGAGATCCAGGAGTTCTTCGCCGCGCGCGAG 796
Db 588 GCCACGCGCGCGAAGATCCGCGCGCGCGCGAGATCGAGGCGTTCTTCGCGCGCGCGCGAG 529
QY 797 GCGGCCAGGCGCGGCTTTGCTTCCAAAGTACAACTTCGACTTCGTCGCGGCGTGCCC 856
Db 528 GAGGCTGAGGCCAAGCGCTTCGCCCGCAAGTACAACTTCGACGTCGTTCCGCGCGTGCCC 469
QY 857 CTCGACGCGCGCGCGGTCGAGTGGCGCGCGGTCGAGTCTGAGC--GAGCGTG 914
Db 468 CTCGACGC---CGGTCGGTTCGAGTGGAATCCCGTGCTCAGCAGCCGGAAGCTGAAGCGAG 412
QY 915 CGTCCGGTGCAAGGTGAAGCTAGAAAG 941
Db 411 CGTGCAGATTAAGCGAAGCTAGAAAG 385

RESULT 9

US-10-333-006-9
; Sequence 9, Application US/10333006
; Publication No. US20040019926A1
; GENERAL INFORMATION:
; APPLICANT: Frankard, Valerie Marie-No. US20040019926A111e S.
; APPLICANT: Peres Bota, Adrian Marius
; APPLICANT: Droual, Anne-Marie
; APPLICANT: Mironov, Vladimir
; APPLICANT: Inz, Dirk
; APPLICANT: Hatzfeld, Yves
; TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
; FILE REFERENCE: 1187-13
; CURRENT APPLICATION NUMBER: US/10/333,006
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: PCT/IB01/01492
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/218,471
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/241,219
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1242
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-333-006-9

Query Match 21.8%; Score 299; DB 16; Length 1242;
Best Local Similarity 68.0%; Pred. No. 7.7e-71;
Matches 577; Conservative 0; Mismatches 210; Indels 62; Gaps 9;

QY 131 CAGATGGGGAAGTACATGCGCAAGTGCAGGGGCGCGCGAGGTCGCGCCCGTC 190
Db 15 CAGATGGGCAAGTACATGAGAAAGTTAGGGGGGCGCACGGGGAGGAGTTGGCCGCCATG 74
QY 191 GAGGTTACGCAAGTCTCGGCGCTCCGGACGAGGTCAGGTCAGGTCGCGGGCGGACCGCGG-- 248
Db 75 GAGGTCACGCAAGTGGTGGCGTCCGGACGAGGTCAGGTCGAGGTCGCGAGCGCGCGCGCG 134
QY 249 -----GTGTGCGGAAGTTCGCCCGCGAGGAGGAAAGAGGGCGCGCGGGG 292
Db 135 ACGACGACGAAGGTGCAGGCGGCGTCCGCGCGCTCCACAGGAGGAGGAAGGCGCTGCTG 194
QY 293 GAGCCTGCTGCCGCGCTGAGCGCTGGTGGGACGCGGAAAGCTGTACATCCACCTGCGT 352
Db 195 CCGACGCGGTCGTCGGGACTACTCGCCGTGACCGCGGAGCTGCTACCTCCAGCTGAGG 254
QY 353 AGCCGCTGCTGTCATGGCACCGCTCAGCCGCGAGCGCGCTGCTGCTGCTGCTGCTGCTG 412
Db 255 AGCCGCTGCTGTCATGGCGCGCGCT---GAGGCGCGCGCGCGCGCGGCTCCGCTT 311
QY 413 CCGTGGAGGCTGCTGATGGCGCTGCAGGACAGCAGGCGCGCGCGCTCGCGCGCGGGCTC 472
Db 312 GTAGCGGAGGCGCGGTTCCGGGAACGAGCGCGCGCGCATGCG---GCGGCTGCGCTC 368
QY 473 TCGGCTGCTCCAGACGCGCTGTCGCTGAGTGAATTTGGGCTTGGGGGTCAGCGCGGAGC 532
Db 369 TCGGCTGCTCCAGACGCGCTGCTC-----CGTGAACGCGCGCGCTCAGGACAGGAGC 422
QY 533 CACACCTGCCGCTCCTACGACGCTGCAGAGGTCGCGGGGATCACGTCCTGCTGCTGCTG 592
Db 423 CTCGCTGCCGCTC---CGACGTCGCGGAGGCGAGGCGAGCATGTCGCCGAGG--- 474
QY 593 TCGGCGGCGAGCAACTCCGGGAGCGCGCGCGAGCGCGGAGGCGGAGAGAGCGCCATCG 652
Db 475 -GCTCCGCGAGCGACTCGCGGAGCGCGCGTACCGCGAGGAGAGAAACAACACTCCATCA 533
QY 653 AGCCGGGCGCACGCGAGCTCAGCGATCTGAGTCTGAGTCTGCGGGGCGCAAGA----- 707
Db 534 AGCTTTCTCCCCGCGAGGTGAGCGATCTGAGTCTGAGTCTGAGTCTGAGGAGGAGAGCGC 593
QY 708 -----CTGGCCCGTCGCTACCGCGCGCAACGCGCGCTGCGGAGCTGATCGTG 754
Db 594 AGCCGTCCACTACTTCTCGGGAACAGCCTCAGCACAGCAAGCCACGCGCGCGAAGATT 653
QY 755 CCGCCAGCACACGAGATCCAGGAGTTCTTCGCCCGCGCGCGAGGCGCGCCAGGCGCGC 814
Db 654 CCGCCGCGCGCGCGAGATCGAGGCGTTCTTCGCGCGCGCGCGAGGAGGCTGAGGCGCGC 713
QY 815 TTTGCTTCCAAAGTACAACTTCGACTTCGTCGCGCGCGCTGCCCTCGACGCGCGCGCGCG 874
Db 714 TTCGCCCGCAAGTACAACTTCGACGCTGTTGCGGCGCTGCCCTCGACGC---CGGTCGG 770
QY 875 TTCGAGTGGCGCGGCTGGTGGTGCATCTGAAGC--GAGCGTGGCTCCGCTGCAAGGTGAA 932
Db 771 TTCGAGTGGACTCCGCTGGTGCAGCAGCCGGAAGCTGAAGCGGAGCGCTGCAGATTAAAG 830
QY 933 GCTAGAAAG 941
Db 831 GCTAGAAAG 839

RESULT 10

US-10-767-701-24179
; Sequence 24179, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701

```

; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 24179
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 30
US-10-767-701-24179

```

| Query Match | 19.9% | Score 272.4 | DB 17 | Length 510 |
|-----------------------|----------------|--|-----------|------------|
| Best Local Similarity | 78.3% | Pred. No. 8.8e-64 | | |
| Matches 407 | Conservative 0 | Mismatches 71 | Indels 42 | Gaps 5 |
| QY | 12 | GGACGCAAGCGGCTGCAGGCAGCAGCGCGCGCAGGCGTTGTGGCCTGTGGGAGAGGAA | 71 | |
| DB | 5 | GAATCCAAAGCGCAAGCGGTCTGCGCAGGCAGCGCGCGCAGGCGTGGGCTGCGAGAG | 64 | |
| QY | 72 | AAGAGAAAGAGGAACCGGCCAAGACAAGCAAGCAGAGGCCAGGGCCGCGGCTTGCCTC | 131 | |
| DB | 65 | AGGAAAAGAGGAACCGGCCAAGAGAAAGCAAGCCGCGGGCG-----GCGGCTTCCGTC | 118 | |
| QY | 132 | AGATGGGAAGTACATGC---GCAAGTGCAGGGCGCGCAGGCGCGGAGTCCGCGCGG | 188 | |
| DB | 119 | AGATGGGAAGTACATGCGCAGCAAGTGCAGGGGGCGCGGGCGCAGGAGTCCGCGCAG | 178 | |
| QY | 189 | TCGAGGTTACGCAGGTCGTGGCGTCCGGACGAGGTCCAGTCCGCGCGCGGCGACCGCG | 248 | |
| DB | 179 | TCGAGGTCACGCAGGTTGTGGCGTCAGGACGAGGTCCAGTCCGCGCGCGGCGACCGCG | 238 | |
| QY | 249 | GTGTCGCGAAGGTCGCCCGCAGGAGGAAGGGCGC-----CGGCGG | 290 | |
| DB | 239 | GTGTCGCGAAGTTCGTCGAGGAGGAAGAAGCGCTGACGCCCGCGCGAACGTGCCG | 298 | |
| QY | 291 | GGGAGCCTGCTGCCCGCGTGAGCGCTGGTGGGACGGCGGAAGCTGTACATCCACCTGC | 350 | |
| DB | 299 | CGGGGGAGCCTGCCCGCGTGGCGCTGGCGGGGACGGCGGGAGCTGTACATCCACCTGC | 358 | |
| QY | 351 | GTAGCCGCATGCTGTTATGGCACCGCCTCAGCCGCGAGCGTGGTTGACTCGGTTCCGA | 410 | |
| DB | 359 | GAAGCCGCATGCTGTTATGGCACCGCCTCAGCCGCGAGCCG-----TCGGTTCCGA | 409 | |
| QY | 411 | CCCCGGTGGAGGCTGCTGATGGCGTGCAGGACAGCAGGGCGCGCGCTCGCGGCGGGC | 470 | |
| DB | 410 | CGCCCGGAGGCTGCT-----GGTGCAGGACAGCAGGGCGCGCGCTCGTGGCTGGC | 463 | |
| QY | 471 | TCTCGGTTGCTCCAGCACGGCGTCGTCCGTGAACTTGGG | 510 | |
| DB | 464 | TCTCGGTTGCTCCAGCACGGCGTCGTCCGTGGACGTCGG | 503 | |

RESULT 11
US-10-333-006-5
; Sequence 5, Application US/10333006
; Publication No. US20040019926A1
; GENERAL INFORMATION:
; APPLICANT: Frankard, Valerie Marie-No. US20040019926A1lle S.
; APPLICANT: Peres Bota, Adrian Marius
; APPLICANT: Droual, Anne-Marie
; APPLICANT: Mironov, Vladimir
; APPLICANT: Inz,, Dirk
; APPLICANT: Hatzfeld, Yves
; TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
; FILE REFERENCE: 1187-13
; CURRENT APPLICATION NUMBER: US/10/333,006
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: PCT/IB01/01492
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/218,471
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/241,219
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 59

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (32)..(32)
; OTHER INFORMATION: n = A, T, C or G
US-10-333-006-5

```

| | Query Match | 16.0%; | Score 219; | DB 16; | Length 568; |
|----|-----------------------|--|--------------------|----------------|-------------------|
| | Best Local Similarity | 84.6%; | Pred. No. 3.5e-49; | | |
| | Matches 335; | Conservative | 0; | Mismatches 36; | Indels 25; Gaps 7 |
| QY | 563 | GCTGTGGGGGATCACGTCTTGGTGGATGTCTCGGCGGCGAGCAACTCGGGAGCGGCCCA | 622 | | |
| Db | 2 | GGTTACGGGGATCACGTCC--CGGATGTCGTCNCCGCGAGCAACTCGGGAGCGTCCCG | 58 | | |
| QY | 623 | GACCGGAGAGGCGAGACGACGCGCA--TCGAGCCGGGCGCA---CGGCGAGCTCAGC | 676 | | |
| Db | 59 | GACCGGAGAGGAGAGACGACGCCATCGTCGAGCCGGGCGACGGCGGCGAGCTCAGC | 118 | | |
| QY | 677 | GATCTGGAGTCGGATCTGGCGG--GGCACAAAGACTGGC-----CCGTGCTACCGGCG | 727 | | |
| Db | 119 | GATCTGGAGTCGGATCTGGTGGGCGGCAGAAAGACTGGCTGCTCGTCGTCGCGGCGACA | 178 | | |
| QY | 728 | GCAACGCCGGCTGCGGAGTGATCGTGCCGCCAGCACACGAGATCCAGGAGTTCTTCGCC | 787 | | |
| Db | 179 | ACAACATCGGCTCGGAGCTGATCGTGCCGCCAGCACAGGAGATCCAGGAATCTTCGCG | 238 | | |
| QY | 788 | GCCGCCGAGGGGCCCGCCAGCGCTTTGCTTCCAAAGTACAACCTTCGACTTCGTCGCG | 847 | | |
| Db | 239 | GCCGCCGAGGGGCCCATGTGCCAACGCTTTGCTTCCAAAGTACAACCTTCGACTTCGTCGCG | 298 | | |
| QY | 848 | GGCGTGCCCTCGACGCCGGCGCGCCGGTTCGAGTGGGGCGCGGTGGTCAGCATCTGAAGC | 907 | | |
| Db | 299 | GGCGTGCCCTCGACGCG--CGGCCGGTTCGAGTGGACGCCAGGGGTCAGCATCTGAAGC | 355 | | |
| QY | 908 | GAGCGTGCGTCCGGTGCAAGGTGAAGCTAGAAAGAG | 943 | | |
| Db | 356 | GAGCGTG----CGGTGCAAGGTGAAGCTACTACTAG | 387 | | |

RESULT 12
US-10-653-595-104
; Sequence 104, Application US/10653595
; Publication No. US20040048304A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et. al.
; TITLE OF INVENTION: 95 Human secreted proteins
; FILE REFERENCE: PZ027P1C1
; CURRENT APPLICATION NUMBER: US/10/653,595
; CURRENT FILING DATE: 2003-09-03
; PRIOR APPLICATION NUMBER: US 09/397945
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: PCT/US99/05804
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/078,566
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,576
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,573
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,574
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,579
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/080,314
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080,312
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/078,578


```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3166)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (3171)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (3264)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (3308)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (3309)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-846-83

Query Match      8.1%; Score 110.8; DB 9; Length 3313;
Best Local Similarity 89.4%; Pred. No. 2.9e-19;
Matches 118; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1240 AAAGAAAGATCTGTTTGTCTCAAAAAAAAAAAAAAAAAAGGCGCGCCGC 1299
Db 3107 AATAAAACATTTTGGATTTTCAAAAAAAAAAAAAAAAAAGGCGCGCCGN 3166

QY 1300 TCTAGAGGATCCAGCTTACGTACGCGTGCGATGCGACGTCTCTTCTATAGTGTC 1359
Db 3167 TCTANAGGATCCAGCTTACGTACGCGTGCGATGCGACGTCTCTTCTATAGTGTC 3226

QY 1360 CCTAAATTCAAT 1371
Db 3227 CCTAAATTCAAT 3238
```

```
RESULT 15
US-10-091-483-83
; Sequence 83, Application US/10091483
; Publication No. US20030049650A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ12C1
; CURRENT APPLICATION NUMBER: US/10/091,483
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 348
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 83
; LENGTH: 3313
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3166)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3171)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3264)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3308)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3309)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-091-483-83
```

Query Match 8.1%; Score 110.8; DB 15; Length 3313;

```
Best Local Similarity 89.4%; Pred. No. 2.9e-19;
Matches 118; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1240 AAAGAAAGATCTGTTTGTCTCAAAAAAAAAAAAAAAAAAGGCGCGCCGC 1299
Db 3107 AATAAAACATTTTGGATTTTCAAAAAAAAAAAAAAAAAAGGCGCGCCGN 3166

QY 1300 TCTAGAGGATCCAGCTTACGTACGCGTGCGATGCGACGTCTCTTCTATAGTGTC 1359
Db 3167 TCTANAGGATCCAGCTTACGTACGCGTGCGATGCGACGTCTCTTCTATAGTGTC 3226

QY 1360 CCTAAATTCAAT 1371
Db 3227 CCTAAATTCAAT 3238
```

Search completed: October 2, 2004, 06:02:18
Job time : 713 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2004, 05:36:42 ; Search time 3331 Seconds
(without alignments)
12299.883 Million cell updates/sec

Title: US-09-993-808B-1
Perfect score: 1372
Sequence: 1 cccacgcgtcggacgcaag.....agtgtaacacctaattcattc 1372

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 389 | 28.4 | 526 | 14 | CD444277 EL01N0438 |
| 2 | 197 | 14.4 | 890 | 29 | CG218373 OG1DJ10TV |
| 3 | 197 | 14.4 | 1010 | 28 | CC390569 PUHBT76TB |
| 4 | 193 | 14.1 | 753 | 29 | CC729673 OGUBY35TH |

| | | | | | |
|----|-----|------|-----|----|----------|
| 5 | 193 | 14.1 | 791 | 29 | CG324885 |
| 6 | 193 | 14.1 | 825 | 29 | CG265472 |
| 7 | 193 | 14.1 | 917 | 29 | CG265481 |
| 8 | 178 | 13.0 | 547 | 28 | BH777311 |
| 9 | 167 | 12.2 | 884 | 29 | CG223284 |
| 10 | 167 | 12.2 | 885 | 29 | CG324875 |
| 11 | 167 | 12.2 | 887 | 29 | CG223296 |
| 12 | 160 | 11.7 | 707 | 11 | AY110319 |
| 13 | 152 | 11.1 | 986 | 28 | CC002416 |
| 14 | 152 | 11.1 | 987 | 28 | CC002419 |
| 15 | 143 | 10.4 | 645 | 29 | CG136784 |
| 16 | 131 | 9.5 | 225 | 14 | CD957606 |
| 17 | 119 | 8.7 | 505 | 14 | CA402478 |
| 18 | 112 | 8.2 | 741 | 29 | CC683397 |
| 19 | 112 | 8.2 | 790 | 28 | CC390573 |
| 20 | 112 | 8.2 | 835 | 29 | CG218361 |
| 21 | 109 | 7.9 | 669 | 12 | BG924526 |
| 22 | 108 | 7.9 | 172 | 14 | CF643461 |
| 23 | 108 | 7.9 | 397 | 12 | BG898272 |
| 24 | 108 | 7.9 | 402 | 12 | BG897464 |
| 25 | 108 | 7.9 | 404 | 12 | BG896735 |
| 26 | 108 | 7.9 | 433 | 12 | BG898370 |
| 27 | 108 | 7.9 | 441 | 12 | BG927036 |
| 28 | 108 | 7.9 | 473 | 12 | BG898628 |
| 29 | 108 | 7.9 | 480 | 12 | BG928051 |
| 30 | 108 | 7.9 | 559 | 12 | BG924503 |
| 31 | 108 | 7.9 | 583 | 12 | BG927127 |
| 32 | 108 | 7.9 | 586 | 12 | BG926479 |
| 33 | 108 | 7.9 | 591 | 12 | BG928233 |
| 34 | 108 | 7.9 | 595 | 12 | BG924343 |
| 35 | 108 | 7.9 | 608 | 12 | BG896941 |
| 36 | 108 | 7.9 | 609 | 12 | BG924870 |
| 37 | 108 | 7.9 | 609 | 12 | BG929210 |
| 38 | 108 | 7.9 | 618 | 12 | BG928558 |
| 39 | 108 | 7.9 | 620 | 12 | BG925002 |
| 40 | 108 | 7.9 | 628 | 12 | BG928414 |
| 41 | 108 | 7.9 | 631 | 12 | BG927739 |
| 42 | 108 | 7.9 | 632 | 12 | BG926073 |
| 43 | 108 | 7.9 | 633 | 12 | BG899001 |
| 44 | 108 | 7.9 | 634 | 12 | BG899738 |
| 45 | 108 | 7.9 | 635 | 12 | BG924853 |

ALIGNMENTS

RESULT 1
CD444277
LOCUS CD444277 526 bp mRNA linear EST 03-JUN-2003
DEFINITION EL01N0438A02.b Endosperm_4 Zea mays cDNA, mRNA sequence.
ACCESSION CD444277
VERSION CD444277.1 GI:31359920
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 526)
AUTHORS Lai, J., Dey, N., Kim, C.S., Becraft, P., Larkins, B., Linton, E. and
Messing, J.
TITLE Sequencing of the maize endosperm ESTs
JOURNAL Unpublished (2002)
COMMENT Contact: Lai, Jinsheng
Dr. Joachim Messing's lab
Waksman Institute, Rutgers University
190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
Tel: 732-445-3801
Fax: 732-445-5735
Email: jlai@waksman.rutgers.edu
Seq primer: T3
Location/Qualifiers
source 1. 526


```
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 632 AGCGAGAGACGACGCCATCGAGCCGGCGGCACGGCGAGCTCAGCGATCTGGAGTCGGAT 691
    |||||
Db 177 AGCGAGAGACGACGCCATCGAGCCGGCGGCACGGCGAGCTCAGCGATCTGGAGTCGGAT 236

QY 692 CTGGCGGGGCACAAGACTGGCCCGTCGCTACCGGGGCAACGCCGGCTGCGGAGCTGATC 751
    |||||
Db 237 CTGGCGGGGCACAAGACTGGCCCGTCGCTACCGGGGCAACGCCGGCTGCGGAGCTGATC 296

QY 752 GTGCCGCCAGCACACGAGATCCAGGAGTTCTTCGCCCGCCGCCGAGGGCCACAGGCCAAG 811
    |||||
Db 297 GTGCCGCCAGCACACGAGATCCAGGAGTTCTTCGCCCGCCGCCGAGGGCCACAGGCCAAG 356

QY 812 CGCTTTGCTTCCAAGTA 828
    |||||
Db 357 CGCTTTGCTTCCAAGTA 373

RESULT 4
CC729673/c
LOCUS
DEFINITION
  OGU8Y35TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0404F21,
  genomic survey sequence.
ACCESSION
  CC729673
VERSION
  CC729673.1 GI:32148606
KEYWORDS
  GSS.
SOURCE
  Zea mays
ORGANISM
  Zea mays
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
  clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
  1 (bases 1 to 753)
AUTHORS
  Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
  Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
  Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  Consortium for Maize Genomics
  Unpublished (2002)
  Other_GSSs: OGU8Y35TV
  Contact: Cathy Whitelaw
TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@tigr.org
  Seq primer: TR
  Class: sheared ends.
FEATURES
  source
    Location/Qualifiers
      1..753
        /organism="Zea mays"
        /mol_type="genomic DNA"
        /strain="B73"
        /db_xref="taxon:4577"
        /clone="ZMMBMA0404F21"
        /clone_lib="ZM 0.7 1.5_KB"
        /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
        methylation filtered genomic DNA library"
ORIGIN
  Query Match 14.1%; Score 193; DB 29; Length 753;
  Best Local Similarity 100.0%; Pred. No. 9.3e-29;
  Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 636 GAGAGACGACGCCATCGAGCCGGCGGCACGGCGAGCTCAGCGATCTGGAGTCGGATCTGG 695
    |||||
Db 406 GAGAGACGACGCCATCGAGCCGGCGGCACGGCGAGCTCAGCGATCTGGAGTCGGATCTGG 347

QY 696 CGGGGCACAAGACTGGCCCGTCGCTACCGGGGCAACGCCGGCTGCGGAGCTGATCGTGC 755
    |||||
Db 346 CGGGGCACAAGACTGGCCCGTCGCTACCGGGGCAACGCCGGCTGCGGAGCTGATCGTGC 287

QY 756 CGCCAGCACACGAGATCCAGGAGTTCTTCGCCCGCCGCCGAGGGCCCGCCAGCGCGCT 815
    |||||
```

```
Db 286 CGCCAGCACACGAGATCCAGGAGTTCTTCGCCCGCCGCCGAGGGCCCGCCAGCGCGCT 227

QY 816 TTGCTTCCAAGTA 828
    |||||
Db 226 TTGCTTCCAAGTA 214

RESULT 5
CG324885
LOCUS
DEFINITION
  OGXDN04TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0661B08,
  genomic survey sequence.
ACCESSION
  CG324885
VERSION
  CG324885.1 GI:34242151
KEYWORDS
  GSS.
SOURCE
  Zea mays
ORGANISM
  Zea mays
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
  clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
  1 (bases 1 to 791)
AUTHORS
  Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
  Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
  Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  Consortium for Maize Genomics
  Unpublished (2002)
  Other_GSSs: OGXDN04TH
  Contact: Cathy Whitelaw
TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@tigr.org
  Seq primer: TF
  Class: sheared ends.
FEATURES
  source
    Location/Qualifiers
      1..791
        /organism="Zea mays"
        /mol_type="genomic DNA"
        /strain="B73"
        /db_xref="taxon:4577"
        /clone="ZMMBMA0661B08"
        /clone_lib="ZM 0.7 1.5_KB"
        /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
        methylation filtered genomic DNA library"
ORIGIN
  Query Match 14.1%; Score 193; DB 29; Length 791;
  Best Local Similarity 100.0%; Pred. No. 8.9e-29;
  Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 636 GAGAGACGACGCCATCGAGCCGGCGGCACGGCGAGCTCAGCGATCTGGAGTCGGATCTGG 695
    |||||
Db 200 GAGAGACGACGCCATCGAGCCGGCGGCACGGCGAGCTCAGCGATCTGGAGTCGGATCTGG 259

QY 696 CGGGGCACAAGACTGGCCCGTCGCTACCGGGGCAACGCCGGCTGCGGAGCTGATCGTGC 755
    |||||
Db 260 CGGGGCACAAGACTGGCCCGTCGCTACCGGGGCAACGCCGGCTGCGGAGCTGATCGTGC 319

QY 756 CGCCAGCACACGAGATCCAGGAGTTCTTCGCCCGCCGCCGAGGGCCCGCCAGCGCGCT 815
    |||||
Db 320 CGCCAGCACACGAGATCCAGGAGTTCTTCGCCCGCCGCCGAGGGCCCGCCAGCGCGCT 379

QY 816 TTGCTTCCAAGTA 828
    |||||
Db 380 TTGCTTCCAAGTA 392

RESULT 6
CG265472/c
LOCUS
DEFINITION
  OGU2CM60TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0764I23,
  genomic survey sequence.
  825 bp DNA linear GSS 25-AUG-2003
```

ACCESSION CG265472
VERSION CG265472.1 GI:34177613
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 825)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfling,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OG2CM60TV
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1. .825
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBMA0764I23"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
FEATURES
source
Query Match 14.1%; Score 193; DB 29; Length 825;
Best Local Similarity 100.0%; Pred. No. 8.5e-29;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
ORIGIN
636 GAGAGACGACGCCATCGAGCGGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGATCTGG 695
|||||
816 GAGAGACGACGCCATCGAGCGGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGATCTGG 757
696 CGGGGCACACAGACTGGCCCGTCGCTACCGGGCGCAACCGCGGCTGCGGAGCTGATCGTGC 755
|||||
756 CGGGGCACACAGACTGGCCCGTCGCTACCGGGCGCAACCGCGGCTGCGGAGCTGATCGTGC 697
756 CGCCAGCACACGAGATCAGGAGTTCTTCGCCCGCGGAGCGGCCCGCCAGGCCAAGCGCT 815
|||||
696 CGCCAGCACACGAGATCAGGAGTTCTTCGCCCGCGGAGCGGCCCGCCAGGCCAAGCGCT 637
816 TTGCTTCCAAGTA 828
|||||
636 TTGCTTCCAAGTA 624
RESULT 7
CG265481
LOCUS CG265481 917 bp DNA linear GSS 25-AUG-2003
DEFINITION CG2CM60TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0764I23,
genomic survey sequence.
ACCESSION CG265481
VERSION CG265481.1 GI:34177622
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 917)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfling,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.

TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other_GSSs: OG2CM60TH
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1. .917
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBMA0764I23"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
FEATURES
source
Query Match 14.1%; Score 193; DB 29; Length 917;
Best Local Similarity 100.0%; Pred. No. 7.7e-29;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
ORIGIN
636 GAGAGACGACGCCATCGAGCGGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGATCTGG 695
|||||
266 GAGAGACGACGCCATCGAGCGGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGATCTGG 325
696 CGGGGCACACAGACTGGCCCGTCGCTACCGGGCGCAACCGCGGCTGCGGAGCTGATCGTGC 755
|||||
326 CGGGGCACACAGACTGGCCCGTCGCTACCGGGCGCAACCGCGGCTGCGGAGCTGATCGTGC 385
756 CGCCAGCACACGAGATCAGGAGTTCTTCGCCCGCGGAGCGGCCCGCCAGGCCAAGCGCT 815
|||||
386 CGCCAGCACACGAGATCAGGAGTTCTTCGCCCGCGGAGCGGCCCGCCAGGCCAAGCGCT 445
816 TTGCTTCCAAGTA 828
|||||
446 TTGCTTCCAAGTA 458
RESULT 8
BH777311/c
LOCUS BH777311 547 bp DNA linear GSS 28-MAR-2002
DEFINITION fzmb013f023f06f0 fzmb filtered library Zea mays genomic clone
fzmb013f023f06 5', genomic survey sequence.
ACCESSION BH777311
VERSION BH777311.1 GI:19779747
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 547)
AUTHORS Budiman,M.A., Freese,R.G., Bedell,J.A., Nunberg,A.N. and Lakey,N.D.
TITLE GeneThresher methylation filtered genomic sequences from maize
JOURNAL Unpublished (2002)
COMMENT Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: fzmb013f023 row: f column: 06
Seq primer: M13 forward
Class: shotgun
High quality sequence stop: 547.
Location/Qualifiers
1. .547
FEATURES
source

/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="Mol17"
/db_xref="taxon:4577"
/clone="fzmb013f023f06"
/clone_lib="fzmb filtered library"
/note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells."

ORIGIN

Query Match 13.0%; Score 178; DB 28; Length 547;
Best Local Similarity 100.0%; Pred. No. 1e-25;
Matches 178; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 632 AGCGGAGAGACGACGCCATCGAGCCGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGAT 691
|||||
Db 294 AGCGGAGAGACGACGCCATCGAGCCGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGAT 235
|||||

QY 692 CTGGCGGGGCACAAGACTGCGCCGTCGCTACCGCGGGCAACCGCGGCTGCGGAGCTGATC 751
|||||
Db 234 CTGGCGGGGCACAAGACTGCGCCGTCGCTACCGCGGGCAACCGCGGCTGCGGAGCTGATC 175
|||||

QY 752 GTGCCGCCAGCACACGAGATCCAGGAGTTCTTCCGCCGCGCGAGCGGCCCGCCAGGCCA 809
|||||
Db 174 GTGCCGCCAGCACACGAGATCCAGGAGTTCTTCCGCCGCGCGAGCGGCCCGCCAGGCCA 117
|||||

RESULT 9

CG223284/c
LOCUS
DEFINITION
CGWAU17TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma0524C09,
genomic survey sequence.

ACCESSION
CG223284
VERSION
CG223284.1 GI:34123172
KEYWORDS
GSS.

SOURCE
Zea mays
ORGANISM
Zea mays

REFERENCE

AUTHORS
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 884)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

TITLE
Consortium for Maize Genomics
JOURNAL
Unpublished (2002)
COMMENT
Other GSSs: OGWAU17TV
Contact: Cathy Whitelaw

FEATURES

source
Location/Qualifiers
1..884
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0524C09"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 12.2%; Score 167; DB 29; Length 884;
Best Local Similarity 100.0%; Pred. No. 8.5e-24;
Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 999 AACAAATTAAGCAGCTTTATATAGCCTAAGCTAACCAACCACCAATTCATCTCGTCCAAATGC 1058
|||||
Db 867 AACAAATTAAGCAGCTTTATATAGCCTAAGCTAACCAACCACCAATTCATCTCGTCCAAATGC 808
|||||

QY 1059 ATGCCTTGCTTTTCTCTGGAGCTAGCAGGAGCGTAGTTATTATTAGTACTACTTTACTT 1118
|||||
Db 807 ATGCCTTGCTTTTCTCTGGAGCTAGCAGGAGCGTAGTTATTATTAGTACTACTTTACTT 748
|||||

QY 1119 ATTCAGAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTA 1165
|||||
Db 747 ATTCAGAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTA 701
|||||

RESULT 10

CG324875/c
LOCUS
DEFINITION
CG324875 Zea mays genomic clone ZMMBma0661B08,
genomic survey sequence.

ACCESSION
CG324875
VERSION
CG324875.1 GI:34242141
KEYWORDS
GSS.

SOURCE
Zea mays
ORGANISM
Zea mays

REFERENCE

AUTHORS
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 885)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

TITLE
Consortium for Maize Genomics
JOURNAL
Unpublished (2002)
COMMENT
Other GSSs: OGXDN04TV
Contact: Cathy Whitelaw

FEATURES

source
Location/Qualifiers
1..885
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0661B08"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 12.2%; Score 167; DB 29; Length 885;
Best Local Similarity 100.0%; Pred. No. 8.5e-24;
Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 999 AACAAATTAAGCAGCTTTATATAGCCTAAGCTAACCAACCACCAATTCATCTCGTCCAAATGC 1058
|||||
Db 414 AACAAATTAAGCAGCTTTATATAGCCTAAGCTAACCAACCACCAATTCATCTCGTCCAAATGC 355
|||||

QY 1059 ATGCCTTGCTTTTCTCTGGAGCTAGCAGGAGCGTAGTTATTATTAGTACTACTTTACTT 1118
|||||
Db 354 ATGCCTTGCTTTTCTCTGGAGCTAGCAGGAGCGTAGTTATTATTAGTACTACTTTACTT 295
|||||

QY 1119 ATTCAGAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTA 1165
|||||
Db 294 ATTCAGAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGTA 248
|||||

RESULT 11

CG223296

LOCUS CG223296 887 bp DNA linear GSS 22-AUG-2003
DEFINITION OGWAU17TV ZM_0.7_1.5 KB Zea mays genomic clone ZMMBMA0524C09,
genomic survey sequence.
ACCESSION CG223296
VERSION CG223296.1 GI:34123184
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 887)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
CONSORTIUM for Maize Genomics
UNPUBLISHED (2002)
OTHER GSSs: OGWAU17TH
CONTACT: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES Location/Qualifiers
source 1..887
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0524C09"
/clone_lib="ZM 0.7 1.5 KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 12.2%; Score 167; DB 29; Length 887;
Best Local Similarity 100.0%; Pred. No. 8.5e-24;
Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 999 AACAAATTAAGCAGCTTTATATAGCCTTAAGCTAAGCTAACCCACCACCATTCATCTCGTCCAAATGC 1058
Db 296 AACAAATTAAGCAGCTTTATATAGCCTTAAGCTAAGCTAACCCACCACCATTCATCTCGTCCAAATGC 355
QY 1059 ATGCCTTGCTTTTCTCTGGAGCTAGCAGGAGCGTAGTTATTATTAGTACTACTTTACTT 1118
Db 356 ATGCCTTGCTTTTCTCTGGAGCTAGCAGGAGCGTAGTTATTATTAGTACTACTTTACTT 415
QY 1119 ATTACAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGA 1165
Db 416 ATTACAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGA 462
RESULT 12
LOCUS AY110319 707 bp mRNA linear HTC 17-OCT-2002
DEFINITION Zea mays CL528_-1 mRNA sequence.
ACCESSION AY110319
VERSION AY110319.1 GI:21214635
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 707)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)

LOCUS CG223296 887 bp DNA linear GSS 22-AUG-2003
DEFINITION OGWAU17TV ZM_0.7_1.5 KB Zea mays genomic clone ZMMBMA0524C09,
genomic survey sequence.
ACCESSION CG223296
VERSION CG223296.1 GI:34123184
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 887)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
CONSORTIUM for Maize Genomics
UNPUBLISHED (2002)
OTHER GSSs: OGWAU17TH
CONTACT: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES Location/Qualifiers
source 1..887
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0524C09"
/clone_lib="ZM 0.7 1.5 KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 12.2%; Score 167; DB 29; Length 887;
Best Local Similarity 100.0%; Pred. No. 8.5e-24;
Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 999 AACAAATTAAGCAGCTTTATATAGCCTTAAGCTAAGCTAACCCACCACCATTCATCTCGTCCAAATGC 1058
Db 296 AACAAATTAAGCAGCTTTATATAGCCTTAAGCTAAGCTAACCCACCACCATTCATCTCGTCCAAATGC 355
QY 1059 ATGCCTTGCTTTTCTCTGGAGCTAGCAGGAGCGTAGTTATTATTAGTACTACTTTACTT 1118
Db 356 ATGCCTTGCTTTTCTCTGGAGCTAGCAGGAGCGTAGTTATTATTAGTACTACTTTACTT 415
QY 1119 ATTACAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGA 1165
Db 416 ATTACAGGTTATCTTGACCCCGATAGATCAATCCGCTTACTGTGA 462
RESULT 12
LOCUS AY110319 707 bp mRNA linear HTC 17-OCT-2002
DEFINITION Zea mays CL528_-1 mRNA sequence.
ACCESSION AY110319
VERSION AY110319.1 GI:21214635
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 707)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)

REFERENCE 2 (bases 1 to 707)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
FEATURES Location/Qualifiers
source 1..707
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:632312"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
ORIGIN
Query Match 11.7%; Score 160; DB 11; Length 707;
Best Local Similarity 100.0%; Pred. No. 2.4e-22;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1006 AAGCAGCTTTATATAGCCTTAAGCTAAGCTAACCCACCACCATTCATCTCGTCCAAATGCATGCCTT 1065
Db 476 AAGCAGCTTTATATAGCCTTAAGCTAAGCTAACCCACCACCATTCATCTCGTCCAAATGCATGCCTT 535
QY 1066 GCTTTTCTCTGGAGCTAGCAGGAGCGTAGTTATTATTAGTACTACTTTACTTTATTCAGA 1125
Db 536 GCTTTTCTCTGGAGCTAGCAGGAGCGTAGTTATTATTAGTACTACTTTACTTTATTCAGA 595
QY 1126 GGTATCTTGACCCCGATAGATCAATCCGCTTACTGTGA 1165
Db 596 GGTATCTTGACCCCGATAGATCAATCCGCTTACTGTGA 635
RESULT 13
LOCUS CC002416 986 bp DNA linear GSS 31-MAR-2003
DEFINITION PUGKB53TB ZM_0.6_1.0 KB Zea mays genomic clone ZMMBTA402I10,
genomic survey sequence.
ACCESSION CC002416
VERSION CC002416.1 GI:29380976
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 986)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other GSSs: PUGKB53TD
CONTACT: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

```
FEATURES
source
Location/Qualifiers
1..986
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa402I10"
/clone_lib="ZM 0.6 1.0 KB"
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

ORIGIN
Query Match 11.1%; Score 152; DB 28; Length 986;
Best Local Similarity 100.0%; Pred. No. 6.1e-21;
Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 632 AGCGGAGAGACGACGCGATCGAGCCGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGAT 691
|
Db 346 AGCGGAGAGACGACGCGATCGAGCCGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGAT 405
|
QY 692 CTGGCGGGGCACAAAGACTGGCCCGTCGCTACCGGGCAACGCCGGCTGCGGAGTGTATC 751
|
Db 406 CTGGCGGGGCACAAAGACTGGCCCGTCGCTACCGGGCAACGCCGGCTGCGGAGTGTATC 465
|
QY 752 GTGCCGCCAGCACACGAGATCCAGGAGTTCTT 783
|
Db 466 GTGCCGCCAGCACACGAGATCCAGGAGTTCTT 497
|

RESULT 14
CC002419/c
LOCUS
DEFINITION PUGKB53TD ZM_0.6_1.0 KB Zea mays genomic clone ZMMBTa402I10,
genomic survey sequence.
ACCESSION CC002419
VERSION CC002419.1 GI:29380979
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 987)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other_GSSs: PUGKB53TB
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
source
Location/Qualifiers
1..987
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa402I10"
/clone_lib="ZM 0.6 1.0 KB"
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

ORIGIN
Query Match 11.1%; Score 152; DB 28; Length 987;
Best Local Similarity 100.0%; Pred. No. 6.1e-21;
Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 632 AGCGGAGAGACGACGCGATCGAGCCGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGAT 691
|
Db 640 AGCGGAGAGACGACGCGATCGAGCCGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGAT 581
|
QY 692 CTGGCGGGGCACAAAGACTGGCCCGTCGCTACCGGGCAACGCCGGCTGCGGAGTGTATC 751
|
Db 580 CTGGCGGGGCACAAAGACTGGCCCGTCGCTACCGGGCAACGCCGGCTGCGGAGTGTATC 521
|
QY 752 GTGCCGCCAGCACACGAGATCCAGGAGTTCTT 783
|
Db 520 GTGCCGCCAGCACACGAGATCCAGGAGTTCTT 489
|

RESULT 15
CG136784
LOCUS
DEFINITION PUIDC25TB ZM_0.6_1.0 KB Zea mays genomic clone ZMMBTa0556E01,
genomic survey sequence.
ACCESSION CG136784
VERSION CG136784.1 GI:34026526
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 645)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other_GSSs: PUIDC25TD
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source
Location/Qualifiers
1..645
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa0556E01"
/clone_lib="ZM 0.6 1.0 KB"
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

ORIGIN
Query Match 10.4%; Score 143; DB 29; Length 645;
Best Local Similarity 100.0%; Pred. No. 5e-19;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 632 AGCGGAGAGACGACGCGATCGAGCCGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGAT 691
|
Db 503 AGCGGAGAGACGACGCGATCGAGCCGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGAT 562
|
QY 692 CTGGCGGGGCACAAAGACTGGCCCGTCGCTACCGGGCAACGCCGGCTGCGGAGTGTATC 751
|
Db 563 CTGGCGGGGCACAAAGACTGGCCCGTCGCTACCGGGCAACGCCGGCTGCGGAGTGTATC 622
|
QY 752 GTGCCGCCAGCACACGAGATCCCA 774
|
Db 623 GTGCCGCCAGCACACGAGATCCA 645
|

Search completed: October 2, 2004, 08:35:34
Job time : 3335 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OMnucleic - nucleic search, using sw model

Run on: October 2, 2004, 04:19:11 ; Search time 562 Seconds
(without alignments)
10371.046 Million cell updates/sec

Title: US-09-993-808B-1
Perfect score: 1372
Sequence: 1 cccacgcgtccgcacgaag.....agtgccacctaattcattc 1372

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_29Jan04.*
1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002s.*
7: geneseqn2003as.*
8: geneseqn2003bs.*
9: geneseqn2003cs.*
10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 1372 | 100.0 | 1372 | 7 | Abv74603 Maize CKI |
| 2 | 322 | 23.5 | 723 | 3 | Aaa95281 Corn cycl |
| 3 | 108 | 7.9 | 149 | 4 | Aal02569 Human rep |
| 4 | 108 | 7.9 | 1316 | 3 | Aac59966 Human sec |
| 5 | 108 | 7.9 | 1557 | 3 | Aac99034 Human pan |
| 6 | 108 | 7.9 | 1557 | 5 | Aas00830 Human cdn |
| 7 | 108 | 7.9 | 2116 | 3 | Aac99033 Human pan |
| 8 | 108 | 7.9 | 2116 | 5 | Aas00847 Human cdn |
| 9 | 106 | 7.7 | 701 | 4 | Aah33453 Human col |
| 10 | 105 | 7.7 | 1928 | 6 | Abk69104 DNA encod |
| 11 | 104 | 7.6 | 2152 | 4 | Aad05156 Human sec |
| 12 | 103 | 7.5 | 845 | 6 | Abq54253 Human ova |
| 13 | 103 | 7.5 | 1088 | 4 | Aad13354 Human sec |
| 14 | 99 | 7.2 | 2049 | 6 | Ab190818 Human pol |
| 15 | 98 | 7.1 | 570 | 7 | Abz55039 Aspergill |
| 16 | 98 | 7.1 | 2657 | 3 | Aac69567 Human sec |
| 17 | 98 | 7.1 | 2657 | 7 | Ada39657 Human sec |
| 18 | 98 | 7.1 | 2657 | 9 | Adc73401 Human sec |
| 19 | 97 | 7.1 | 544 | 4 | Aah35303 Human col |
| 20 | 97 | 7.1 | 652 | 6 | Abq55965 Human ova |
| 21 | 97 | 7.1 | 2454 | 3 | Aac69584 Human sec |
| 22 | 97 | 7.1 | 2454 | 7 | Ada40257 Human sec |
| 23 | 97 | 7.1 | 2454 | 9 | Adc73781 Human sec |

| | | | | | | |
|----|----|-----|------|---|----------|--------------------|
| 24 | 96 | 7.0 | 377 | 4 | AAI98573 | Aai98573 Human exc |
| 25 | 96 | 7.0 | 377 | 4 | AAI64026 | Aai64026 Human bla |
| 26 | 96 | 7.0 | 452 | 6 | ABQ55723 | Abq55723 Human ova |
| 27 | 96 | 7.0 | 762 | 6 | ABQ55923 | Abq55923 Human ova |
| 28 | 96 | 7.0 | 894 | 9 | ADB47734 | Adb47734 Novel hum |
| 29 | 96 | 7.0 | 973 | 8 | AAD56374 | Aad56374 Human sec |
| 30 | 96 | 7.0 | 1138 | 3 | AAF16267 | Aaf16267 Human pro |
| 31 | 96 | 7.0 | 1631 | 4 | AAK64510 | Aak64510 Human imm |
| 32 | 96 | 7.0 | 1958 | 4 | AAH35000 | Aah35000 Human col |
| 33 | 96 | 7.0 | 1958 | 6 | ABL90702 | Ab190702 Human pol |
| 34 | 96 | 7.0 | 1958 | 6 | ABQ54929 | Abq54929 Human ova |
| 35 | 96 | 7.0 | 3240 | 4 | AAF75345 | Aaf75345 Human TGF |
| 36 | 95 | 6.9 | 395 | 4 | AAH35297 | Aah35297 Human col |
| 37 | 95 | 6.9 | 498 | 3 | AAC98321 | Aac98321 Human col |
| 38 | 95 | 6.9 | 606 | 6 | ABQ55926 | Abq55926 Human ova |
| 39 | 95 | 6.9 | 611 | 4 | AAK88301 | Aak88301 Human dig |
| 40 | 95 | 6.9 | 611 | 5 | AAS39445 | Aas39445 cDNA enco |
| 41 | 95 | 6.9 | 611 | 8 | ADB32171 | Adb32171 Human nov |
| 42 | 95 | 6.9 | 657 | 6 | ABQ55900 | Abq55900 Human ova |
| 43 | 95 | 6.9 | 661 | 6 | ABQ55924 | Abq55924 Human ova |
| 44 | 95 | 6.9 | 670 | 6 | ABQ54585 | Abq54585 Human ova |
| 45 | 95 | 6.9 | 684 | 3 | AAC98310 | Aac98310 Human col |

ALIGNMENTS

RESULT 1
ABV74603
ID ABV74603 standard; DNA; 1372 BP.
XX

AC ABV74603;

XX 21-FEB-2003 (first entry)

XX Maize CKI_B coding sequence.

KW Maize; cyclin-dependent kinase inhibitor; CKI; CKI_B; plant; corn;
KW crop yield; root size; plant growth; tassel size; ear size;
KW male sterility; endoreduplication; gene; ds.

XX Zea mays.

XX Key Location/Qualifiers
FH CDS 134..904
FT /*tag= a
FT /product= "CKI_B"

XX WO200281623-A2.

XX 17-OCT-2002.

XX 06-NOV-2001; 2001WO-US044038.

XX 07-NOV-2000; 2000US-0246349P.

XX (PION-) PIONEER HI-BRED INT INC.
XX (ARIZ-) ARIZONA BOARD OF REGENTS.

XX Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;

XX WPI; 2003-058511/05.

XX P-PSDB; ABB98757.

XX Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
XX involved in cell cycle regulation, and useful for altering cell cycle
XX protein content, cell cycle progression, cell number and composition of
XX plants.

XX Claim 1; Page 64-65; 69pp; English.

XX The present sequence is the coding sequence (I) for maize cyclin-
XX dependent kinase inhibitor (CKI), CKI_B. (I) is useful for modulating the


```
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
PA
XX Klein TM, Weng Z, Cahoon RE;
PI
XX WPI; 2000-679375/66.
DR
DR P-PSDB; AAB26250.
XX
XX Cyclin dependent kinase inhibitor sequences, useful for identifying
PT herbicides and plant growth inhibitors.
PT
XX Claim 2; Page 43-44; 58pp; English.
PS
XX The present sequence is the coding sequence for the corn cyclin-dependent
CC kinase inhibitor (CDKI). It was isolated by searching a contig comprising
CC cDNA from corn roots and ear leaf sheath for sequences similar to those
CC encoding the CDKI from Chenopodium rubrum, Caenorhabditis elegans and
CC Arabidopsis thaliana. CDKI is involved in the cell cycle, and may promote
CC or inhibit cell division and growth. The coding sequence and the protein
CC it encodes are useful in the production of transgenic plants which
CC produce increased or decreased amounts of the CDKI protein, in the
CC identification of herbicides, in genetic and physical mapping and in the
CC isolation of the CDKI gene in other organisms
XX
SQ Sequence 723 BP; 160 A; 202 C; 210 G; 151 T; 0 U; 0 Other;

Query Match 23.5%; Score 322; DB 3; Length 723;
Best Local Similarity 100.0%; Pred. No. 9.5e-100;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 636 GAGAGACGACGCCATCGAGCCGGCGCACGGCGAGCTACGGATCTGGAGTCGGATCTGG 695
Db |||||
QY 696 CGGGGCACAAGACTGGCCCGTGGCTACCGGGCAACGCCGCTCGGAGTCGTCGTC 755
Db |||||
QY 756 CGCAGCACACGAGATCCAGGAGTTCTTCGCCGCCGCCGAGCGGCCAGCCAGCGCT 815
Db |||||
QY 816 TTGCTTCCAAGTACAACTTCGACTTCGTCGCCGCGCTGCCCTCGACGCCGCCGCCGCT 875
Db |||||
QY 876 TCGAGTGGCGCGCGTGGTTCAGCATCTGAAGCGAGCGTGGTCCGGTGCAGGTGAAGCT 935
Db |||||
QY 936 AGAAAGAGAAAAGATGCCCCC 957
Db |||||
QY 412 AGAAAGAGAAAAGATGCCCCC 433

RESULT 3
AAL02569
ID AAL02569 standard; cDNA; 149 BP.
XX
AC AAL02569;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen cDNA SEQ ID NO: 2570.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ss.
XX
OS Homo sapiens.
XX
XX WO200155320-A2.
PN
XX
PD 02-AUG-2001.
```

```
XX 17-JAN-2001; 2001WO-US001339.
PF
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
```

PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465570/50.
DR P-PSDB; AAM96599.

XX
PT Isolated nucleic acid molecule encoding a reproductive system antigen is used in preventing, treating or ameliorating a medical condition.
XX
PS Claim 1; SEQ ID NO 2570; 1297pp + Sequence Listing; English.
XX
CC The present invention provides the protein and coding sequences of a number of human reproductive system related antigens. These can be used in the prevention and treatment of reproductive system disorders, including cancer. The present sequence is a coding sequence of the invention
XX
SQ Sequence 149 BP; 62 A; 30 C; 27 G; 30 T; 0 U; 0 Other;

Query Match 7.9%; Score 108; DB 4; Length 149;
Best Local Similarity 100.0%; Pred. No. 3.4e-27;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1262 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAGCTTACGT 1321
Db 12 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAGCTTACGT 71

Qy 1322 ACGGTGCATGCGACGTATAGTCTTCTATAGTGTACCTAAATTCA 1369
Db 72 ACGGTGCATGCGACGTATAGTCTTCTATAGTGTACCTAAATTCA 119

RESULT 4
AAC59966
ID AAC59966 standard; cDNA; 1316 BP.
XX
AC AAC59966;
XX
DT 26-JAN-2001 (first entry)
XX
DE Human secreted protein gene 1 SEQ ID NO:11.
XX
KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive; neutropic; neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer; vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic; cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder; wound healing; neurological disease; infectious disease; chromosome identification; ss.
XX
OS Homo sapiens.
XX
PN WO200058356-A1.
XX
PD 05-OCT-2000.
XX
PF 22-MAR-2000; 2000WO-US007535.
XX
PR 26-MAR-1999; 99US-0126511P.
PR 17-DEC-1999; 99US-0172413P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Komatsoulis G;
XX
DR WPI; 2000-594639/56.
DR P-PSDB; AAB34773.
XX
Fifty nucleic acid molecules encoding human secreted proteins, useful in the prevention, treatment and diagnosis of cancer, immune disorders, cardiovascular disorders and neurological diseases.
PT
PT
XX
PS Claim 1; Page 347; 425pp; English.
XX
CC The polynucleotide sequences given in AAC59966 to AAC60015 encode the human secreted proteins given in AAB34773 to AAB34822. AAB34823 to AAB34852 represent human secreted polypeptide sequences and proteins homologous to them, which are given in the exemplification of the present invention. Human secreted proteins have activities based on the tissues

CC and cells the genes are expressed in. Examples of activities include:
CC cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
CC antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
CC vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic; and
CC cardiant. The polynucleotides and polypeptides are useful for preventing,
CC treating or ameliorating a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. The polypeptides
CC can also be used as a food additive or preservative to increase or
CC decrease storage capabilities. The polynucleotides are useful for
CC chromosome identification. They are also useful as probes for diagnosing
CC a disorder related to the female reproductive system, particularly breast
CC and/or ovarian cancer. They are also useful in the gene therapy of breast
CC and ovarian cancer. Nucleic acids, protein, antibodies, agonists and
CC antagonists from the present invention are useful in the diagnosis,
CC treatment and prevention of cancer, immune disorders, cardiovascular
CC disorders, wound healing, neurological diseases and infectious disease.
CC AAC59957 to AAC59965 and AAB34772 represents sequence used in the
CC exemplification of the present invention

SQ Sequence 1316 BP; 377 A; 287 C; 273 G; 372 T; 0 U; 7 Other;

Query Match 7.9%; Score 108; DB 3; Length 1316;
Best Local Similarity 100.0%; Pred. No. 2e-27;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAGCTTACGT 1321
Db 1094 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAGCTTACGT 1153
QY 1322 ACGGTGCGATGCGACGTATAGTCTTCTATAGTGTACCTAAATTCA 1369
Db 1154 ACGGTGCGATGCGACGTATAGTCTTCTATAGTGTACCTAAATTCA 1201

RESULT 5
AAC99034
ID AAC99034 standard; cDNA; 1557 BP.
XX AAC99034;
AC AAC99034;

DT 09-MAR-2001 (first entry)
XX Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:262.
DE Human; pancreas; pancreatic cancer; pancreatic cancer antigen; detection;
XX diagnosis; identification; cytostatic; neuroprotective; nootropic;
KW immunomodulatory; relaxant; contraceptive; gynaecological;
KW antiinflammatory; cardiant; gene therapy; chromosome mapping;
KW linkage analysis; tissue identification; tissue typing; forensic; neural;
KW immune system; muscular; reproductive; gastrointestinal; pulmonary;
KW cardiovascular; renal; proliferative; ss.

XX Homo sapiens.
XX WO200055320-A1.
PD 21-SEP-2000.
XX 08-MAR-2000; 2000WO-US005989.
PF 12-MAR-1999; 99US-0124270P.
XX (HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Ruben SM;
XX WPI; 2000-579444/54.
XX P-PSDB; AAB54269.
XX New nucleic acid that is a pancreatic cancer antigen for preventing,
PT treating, or ameliorating a medical condition, particular pancreatic
PT cancer, or for use in assays for diagnosing a pathological condition.
XX

PS Claim 1; Page 701; 1379pp; English.
XX AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC proteins, called pancreatic cancer antigens, given in AAB54008 to
CC AAB54466. The human pancreatic cancer antigens have cytostatic,
CC neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,
CC gynaecological, cardiant and antiinflammatory activities, and can be used
CC in gene therapy. The polynucleotide and proteins can be used for
CC preventing, treating, or ameliorating a medical condition or in assays
CC for diagnosing a pathological condition or a susceptibility to one in a
CC subject. Binding partners to the proteins and the activity of the
CC proteins can be identified. The pancreatic cancer antigens can be used to
CC detect, treat or prevent pancreatic disorders, especially cancer.
CC Agonists and antagonists to the antigens can be screened for. The
CC pancreatic cancer antigen polynucleotides can be used to design nucleic
CC acid hybridisation probes that can be used in chromosome mapping, linkage
CC analysis, tissue identification and/or typing and a variety of forensic
CC and diagnostic methods. The proteins can be used to generate antibodies
CC which are used to purify, detect and target the polypeptides, including
CC both in vivo and in vitro diagnostic and therapeutic methods. The
CC proteins can be used to treat or prevent neural, immune system, muscular,
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC sequences used in the exemplification of the present invention

SQ Sequence 1557 BP; 412 A; 345 C; 315 G; 480 T; 0 U; 5 Other;

Query Match 7.9%; Score 108; DB 3; Length 1557;
Best Local Similarity 100.0%; Pred. No. 1.9e-27;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAGCTTACGT 1321
Db 1211 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAGCTTACGT 1270
QY 1322 ACGGTGCGATGCGACGTATAGTCTTCTATAGTGTACCTAAATTCA 1369
Db 1271 ACGGTGCGATGCGACGTATAGTCTTCTATAGTGTACCTAAATTCA 1318

RESULT 6
AAS00830
ID AAS00830 standard; cDNA; 1557 BP.

XX AAS00830;
AC AAS00830;
DT 04-JUL-2001 (first entry)
XX Human cDNA clone HSCC19 encoding cancer related protein 4.
DE Human; cancer related protein; HSCC19; food additive; preservative;
KW immunogen; antibody; bone cancer; adrenal cancer; bone marrow cancer;
KW breast cancer; gastrointestinal cancer; liver cancer; lung cancer;
KW urogenital cancer; immune disorder; Addison's disease; allergy;
KW autoimmune haemolytic anaemia; autoimmune thyroiditis; diabetes mellitus;
KW Crohn's disease; multiple sclerosis; rheumatoid arthritis;
KW ulcerative colitis; acquired immunodeficiency syndrome; AIDS;
KW cardiovascular disorder; myocardial ischaemia; wound healing;
KW neurological disorder; Parkinson's disease; Alzheimer's disease;
KW cerebral anoxia; epilepsy; viral infection; bacterial infection;
KW fungal infection; parasitic infection; agonist; antagonist; ss.

XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 259..667
FT /*tag= a
FT /product= "Cancer related protein 4"
FT sig_peptide 259..319
FT /*tag= b
FT mat_peptide 320..664
FT /*tag= c
FT /label= Mature_Cancer_related_protein_4

XX WO200118014-A1.
PN 15-MAR-2001.
XX 30-AUG-2000; 2000WO-US023794.
XX 03-SEP-1999; 99US-0152296P.
PR 06-OCT-1999; 99US-0158003P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Roschke V;
XX WPI; 2001-235186/24.
DR P-PSDB; AAU00868.
XX Twenty nine nucleic acid molecules encoding human cancer associated
PT proteins, useful in the prevention, treatment and diagnosis of cancer,
PT immune disorders, cardiovascular disorders and neurological diseases.
XX Disclosure; Page 385; 427pp; English.
XX The sequence encodes a novel Human cancer related protein. The
CC polynucleotides and polypeptides are useful for preventing, treating or
CC ameliorating a medical condition in e.g. humans, mice, rabbits, goats,
CC horses, cats, dogs, chickens or sheep. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities. The polynucleotide are useful for chromosome
CC identification. The nucleic acids, protein, antibodies, agonists and
CC antagonists are useful in the diagnosis, treatment and prevention of
CC cancer (e.g. cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital), immune disorders
CC (e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,
CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
CC sclerosis, rheumatoid arthritis and ulcerative colitis, acquired
CC immunodeficiency syndrome, AIDS), cardiovascular disorders such as
CC myocardial ischaemias, wound healing, neurological diseases (e.g.
CC Parkinson's disease, Alzheimer's disease, cerebral anoxia and epilepsy)
CC and infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Numerous examples of each type of disorder are given in the
CC specification
XX Sequence 1557 BP; 412 A; 345 C; 315 G; 480 T; 0 U; 5 Other;
SQ Query Match 7.9%; Score 108; DB 5; Length 1557;
Best Local Similarity 100.0%; Pred. No. 1.9e-27;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCAAGCTTACGT 1321
Db 1211 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCAAGCTTACGT 1270
QY 1322 ACGCGTGCATGCGACGTCATAGCTCTTCTATAGTGTCACTAAATTCA 1369
Db 1271 ACGCGTGCATGCGACGTCATAGCTCTTCTATAGTGTCACTAAATTCA 1318
RESULT 7
AAC99033
ID AAC99033 standard; cDNA; 2116 BP.
XX AAC99033;
XX 09-MAR-2001 (first entry)
DT Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:261.
DE Human; pancreas; pancreatic cancer; pancreatic cancer antigen; detection;
XX diagnosis; identification; cytostatic; neuroprotective; nootropic;
KW immunomodulatory; relaxant; contraceptive; gynaecological;
KW antiinflammatory; cardiant; gene therapy; chromosome mapping;
KW linkage analysis; tissue identification; tissue typing; forensic; neural;

KW immune system; muscular; reproductive; gastrointestinal; pulmonary;
KW cardiovascular; renal; proliferative; ss.
XX Homo sapiens.
OS WO200055320-A1.
XX 21-SEP-2000.
XX 08-MAR-2000; 2000WO-US005989.
XX 12-MAR-1999; 99US-0124270P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM;
XX WPI; 2000-579444/54.
DR P-PSDB; AAB54268.
XX New nucleic acid that is a pancreatic cancer antigen for preventing,
PT treating, or ameliorating a medical condition, particular pancreatic
PT cancer, or for use in assays for diagnosing a pathological condition.
XX Claim 1; Page 699-700; 1379pp; English.
PS AAC98773 to AAC99231 encode the human pancreatic cancer associated
XX proteins, called pancreatic cancer antigens, given in AAB54008 to
CC AAB54466. The human pancreatic cancer antigens have cytostatic,
CC neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,
CC gynaecological, cardiant and antiinflammatory activities, and can be used
CC in gene therapy. The polynucleotide and proteins can be used for
CC preventing, treating, or ameliorating a medical condition or in assays
CC for diagnosing a pathological condition or a susceptibility to one in a
CC subject. Binding partners to the proteins and the activity of the
CC proteins can be identified. The pancreatic cancer antigens can be used to
CC detect, treat or prevent pancreatic disorders, especially cancer.
CC Agonists and antagonists to the antigens can be screened for. The
CC pancreatic cancer antigen polynucleotides can be used to design nucleic
CC acid hybridisation probes that can be used in chromosome mapping, linkage
CC analysis, tissue identification and/or typing and a variety of forensic
CC and diagnostic methods. The proteins can be used to generate antibodies
CC which are used to purify, detect and target the polypeptides, including
CC both in vivo and in vitro diagnostic and therapeutic methods. The
CC proteins can be used to treat or prevent neural, immune system, muscular,
CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC sequences used in the exemplification of the present invention
XX Sequence 2116 BP; 606 A; 485 C; 390 G; 629 T; 0 U; 6 Other;
SQ Query Match 7.9%; Score 108; DB 3; Length 2116;
Best Local Similarity 100.0%; Pred. No. 1.8e-27;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCAAGCTTACGT 1321
Db 1977 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCAAGCTTACGT 2036
QY 1322 ACGCGTGCATGCGACGTCATAGCTCTTCTATAGTGTCACTAAATTCA 1369
Db 2037 ACGCGTGCATGCGACGTCATAGCTCTTCTATAGTGTCACTAAATTCA 2084
RESULT 8
AAS00847
ID AAS00847 standard; cDNA; 2116 BP.
XX AAS00847;
AC AAS00847;
XX 04-JUL-2001 (first entry)
DT Human cDNA clone H15BL55 encoding cancer related protein 21.
XX

XX Human; cancer related protein; HISBL55; food additive; preservative;
KW immunogen; antibody; bone cancer; adrenal cancer; bone marrow cancer;
KW breast cancer; gastrointestinal cancer; liver cancer; lung cancer;
KW urogenital cancer; immune disorder; Addison's disease; allergy;
KW autoimmune haemolytic anaemia; autoimmune thyroiditis; diabetes mellitus;
KW Crohn's disease; multiple sclerosis; rheumatoid arthritis;
KW ulcerative colitis; acquired immunodeficiency syndrome; AIDS;
KW cardiovascular disorder; myocardial ischaemia; wound healing;
KW neurological disorder; Parkinson's disease; Alzheimer's disease;
KW cerebral anoxia; epilepsy; viral infection; bacterial infection;
KW fungal infection; parasitic infection; agonist; antagonist; ss.
XX Homo sapiens.
OS
XX
XX
XX Key Location/Qualifiers
FH 342..482
FT /*tag= a
FT /product= "Cancer related protein 21"
XX
XX WO200118014-A1.
XX
XX 15-MAR-2001.
XX
XX 30-AUG-2000; 2000WO-US023794.
XX
XX 03-SEP-1999; 99US-0152296P.
XX 06-OCT-1999; 99US-0158003P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Roschke V;
XX
XX WPI; 2001-235186/24.
XX P-PSDB; AAU00886.
XX
XX Twenty nine nucleic acid molecules encoding human cancer associated
PT proteins, useful in the prevention, treatment and diagnosis of cancer,
PT immune disorders, cardiovascular disorders and neurological diseases.
XX
XX Disclosure; Page 379-380; 427pp; English.
XX
XX The sequence encodes a novel Human cancer related protein. The
CC polynucleotides and polypeptides are useful for preventing, treating or
CC ameliorating a medical condition in e.g. humans, mice, rabbits, goats,
CC horses, cats, dogs, chickens or sheep. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities. The polynucleotide are useful for chromosome
CC identification. The nucleic acids, protein, antibodies, agonists and
CC antagonists are useful in the diagnosis, treatment and prevention of
CC cancer (e.g. cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital), immune disorders
CC (e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,
CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
CC sclerosis, rheumatoid arthritis and ulcerative colitis, acquired
CC immunodeficiency syndrome, AIDS), cardiovascular disorders such as
CC myocardial ischaemias, wound healing, neurological diseases (e.g.
CC Parkinson's disease, Alzheimer's disease, cerebral anoxia and epilepsy)
CC and infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Numerous examples of each type of disorder are given in the
CC specification
XX
XX Sequence 2116 BP; 606 A; 485 C; 390 G; 629 T; 0 U; 6 Other;
SQ
Query Match 7.9%; Score 108; DB 5; Length 2116;
Best Local Similarity 100.0%; Pred. No. 1.8e-27;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1262 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGT 1321
Db 1977 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGT 2036
QY 1322 ACGGTGCATGCGACGTACATAGCTCTTCTATAGTGTACCTAAATTCA 1369

Db 2037 ACGGTGCATGCGACGTACATAGCTCTTCTATAGTGTACCTAAATTCA 2084
RESULT 9
AAH33453
ID AAH33453 standard; cDNA; 701 BP.
XX
AC AAH33453;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:509.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; ss.
XX
OS Homo sapiens.
XX
XX WO200122920-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US026524.
XX
XX 29-SEP-1999; 99US-0157137P.
XX 03-NOV-1999; 99US-0163280P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
PI
XX
XX WPI; 2001-235357/24.
XX P-PSDB; AAG74022.
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers.
XX
XX Claim 1; Page 2588; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
CC proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene therapy
CC and vaccine production. N and P may be used in the prevention, diagnosis
CC and treatment of diseases associated with inappropriate P expression. For
CC example, N and P may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of P by expressing inactive proteins or to
CC supplement the patients own production of P. Additionally, N may be used
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent
CC sequences used in the exemplification of the present invention. N.B.
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
XX to 1052, 7921 and 7922
XX
XX Sequence 701 BP; 200 A; 164 C; 146 G; 181 T; 0 U; 10 Other;
SQ
Query Match 7.7%; Score 106; DB 4; Length 701;
Best Local Similarity 100.0%; Pred. No. 1.1e-26;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1264 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGTAC 1323
Db 459 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGTAC 518
QY 1324 GCGTGCATGCGACGTACATAGCTCTTCTATAGTGTACCTAAATTCA 1369
Db 519 GCGTGCATGCGACGTACATAGCTCTTCTATAGTGTACCTAAATTCA 564

```
RESULT 10
ABK69104
ID ABK69104 standard; cDNA; 1928 BP.
XX AC
XX ABK69104;
DT 02-JUL-2002 (first entry)
XX DE
XX DNA encoding human secreted protein, SEQ ID NO 28.
XX KW Human; secreted protein; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; breast; liver; ischaemia;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;
KW corneal infection; wound healing; ocular disorder; skin aging; sunburn;
KW epithelial cell proliferation; organ transplantation; food additive;
KW food storage; gene; ss.
XX XX
OS Homo sapiens.
XX XX
XX WO200224721-A1.
XX PN
XX 28-MAR-2002.
XX PD
XX 09-JAN-2001; 2001WO-US000544.
XX PF
XX 20-SEP-2000; 2000US-0234211P.
XX PR
XX (HUMA-) HUMAN GENOME SCI INC.
XX PA
XX Komatsoulis GA, Baker KP, Rosen CA, Birse CE, Soppet DR;
XX Olsen HS, Ni J, Fiscella M, Moore PA, Wei P, Ebner R, Duan DR;
XX Shi Y, Choi GH;
XX WPI; 2002-330012/36.
XX DR
XX P-PSDB; AAU96183.
XX DR
XX Isolated nucleic acid molecule encoding a human secreted protein is used
XX in preventing, treating or ameliorating a medical condition.
XX PT
XX Claim 1; Page 464; 562pp; English.
XX PS
XX The invention relates to an isolated nucleic acid molecule (I) encoding a
XX human secreted protein (II). (I) and (II) are used to prevent, treat or
XX ameliorate a medical condition in e.g. humans, mice, rabbits, goats,
XX horses, cats, dogs, chickens or sheep. (I) and (II) are also used in
XX diagnosing a pathological condition or susceptibility to a pathological
XX condition. The antibodies to (II) can also be used in alleviating
XX symptoms associated with the disorders and in diagnostic immunoassays
XX e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA).
XX Disorders which are diagnosed or treated include autoimmune diseases e.g.
XX rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the
XX breast or liver, cardiovascular disorders e.g. cardiac arrest,
XX cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous
XX system disorders e.g. Alzheimer's disease, infections caused by bacteria,
XX viruses and fungi and ocular disorders e.g. corneal infection. The
XX polypeptides can also be used to aid wound healing and epithelial cell
XX proliferation, to prevent skin aging due to sunburn, to maintain organs
XX before transplantation, for supporting cell culture of primary tissues,
XX to regenerate tissues and in chemotaxis. The polypeptides can also be
XX used as a food additive or preservative to increase or decrease storage
XX capabilities. ABK69078-ABK69143 represent human secreted protein coding
XX sequences, PCR primers and related sequences used in cloning and
XX expression of the secreted proteins described in examples of the
XX invention
XX SQ Sequence 1928 BP; 470 A; 525 C; 485 G; 444 T; 0 U; 4 Other;
XX Query Match 7.7%; Score 105; DB 6; Length 1928;
XX Best Local Similarity 100.0%; Pred. No. 1.9e-26;
XX Matches 105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGT 1321
Db 1822 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGT 1881
QY 1322 ACGCGTGCATGCGACGTCATAGCTCTTCTATAGTGTCACTAAAT 1366
Db 1882 ACGCGTGCATGCGACGTCATAGCTCTTCTATAGTGTCACTAAAT 1926
RESULT 11
AAD05156
ID AAD05156 standard; cDNA; 2152 BP.
XX AC
XX AAD05156;
DT 17-JUL-2001 (first entry)
XX DE
XX Human secreted protein-encoding gene 5 cDNA clone HGCAC66, SEQ ID NO:46.
XX KW Human; secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; kidney disorder;
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder; infection; wound healing; vulnery; cell culture;
KW chemotaxis; food additive; gene therapy; binding partner identification;
XX ss.
XX XX
OS Homo sapiens.
XX XX
XX Location/Qualifiers
XX CDS 78..173
XX FT /*tag= a
XX FT /product= "Human secreted protein"
XX sig_peptide 78..140
XX FT /*tag= b
XX mat_peptide 141..170
XX FT /*tag= c
XX FT /product= "Human mature secreted protein"
XX XX
XX WO200134769-A2.
XX PD 17-MAY-2001.
XX XX
XX 01-NOV-2000; 2000WO-US030040.
XX PF
XX 05-NOV-1999; 99US-0163580P.
XX PR 30-JUN-2000; 2000US-0215130P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX XX
XX Ruben SM, Komatsoulis GA, Wei P, Baker KP, Fiscella M;
XX WPI; 2001-308781/32.
XX DR P-PSDB; AAE01267.
XX XX
XX New isolated nucleic acid molecule encoding a human secreted protein is
XX used in preventing, treating or ameliorating a medical condition.
XX PT
XX Claim 1; Page 427-428; 519pp; English.
XX PS
XX AAD05121-AAD05203 represent cDNAs corresponding to 24 human secreted
XX protein genes, and AAE01232-AAE01311 represent the proteins they encode.
XX CC AAE01312-AAE01340 represent human secreted protein variants or fragments.
XX CC The secreted proteins and their genes are useful for preventing, treating
XX or ameliorating medical conditions, e.g., by protein or gene therapy.
XX CC Pathological conditions can be diagnosed by determining the amount of the
XX new protein in a sample or by determining the presence of mutations in
XX the new genes. Specific uses are described for each of the 24 genes,
XX based on the tissues in which they are most highly expressed, and include
```

CC developing products for the diagnosis or treatment of proliferative
CC disorders, cancer, tumours, foetal and developmental abnormalities,
CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,
CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
CC angiogenic disorders, kidney disorders, gastrointestinal disorders,
CC pregnancy-related disorders, endocrine disorders, and infections. The
CC proteins can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues,
CC to regenerate tissues, to identify their cognate ligands or binding
CC partners, and in chemotaxis, and can be used as a food additive or
CC preservative to modify storage properties. Antibodies specific for a
CC protein of the invention can be used in alleviating symptoms associated
CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
CC present sequence represents a human secreted protein-encoding cDNA of the
CC invention

XX Sequence 2152 BP; 605 A; 465 C; 407 G; 664 T; 0 U; 11 Other;

Query Match 7.6%; Score 104; DB 4; Length 2152;
Best Local Similarity 100.0%; Pred. No. 3.9e-26;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1266 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGATCCAAAGCTTACGTACGC 1325
Db |
1869 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGATCCAAAGCTTACGTACGC 1928
QY 1326 GTGCATGCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1369
Db |
1929 GTGCATGCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1972

RESULT 12
ID ABQ54253/c

XX ABQ54253 standard; cDNA; 845 BP.

AC ABQ54253;

XX 22-AUG-2002 (first entry)

XX Human ovarian antigen HCCMD55 cDNA, SEQ ID NO:133.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive; gene; ss.

XX Homo sapiens.

XX WO200200677-A1.

XX 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US018569.

XX 07-JUN-2000; 2000US-0209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-147878/19.

XX P-PSDB; ABP41176.

XX

PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
PT cancer), immune disorders, cardiovascular disorders and neurological
PT diseases.

XX Claim 1; SEQ ID NO 133; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents cDNA encoding a human ovarian antigen of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 845 BP; 182 A; 226 C; 234 G; 199 T; 0 U; 4 Other;

Query Match 7.5%; Score 103; DB 6; Length 845;
Best Local Similarity 100.0%; Pred. No. 1.1e-25;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1267 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGATCCAAAGCTTACGTACGC 1326
Db |
178 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGATCCAAAGCTTACGTACGC 119
QY 1327 TGCATGCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1369
Db |
118 TGCATGCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 76

RESULT 13
AAD13354

ID AAD13354 standard; cDNA; 1088 BP.

XX AAD13354;

XX 16-OCT-2001 (first entry)

XX Human secreted protein-encoding gene 10 cDNA clone HTEGF16, SEQ ID NO:20.

XX Human secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder; infection; wound healing; vulnery; cell culture;
KW chemotaxis; food additive; gene therapy; binding partner identification;

KW ss.
XX Homo sapiens.
OS Location/Qualifiers
XX 160..477
FH /*tag= a
FT /product= "Human secreted protein precursor"
FT 160..282
FT /*tag= b
FT 283..474
FT /*tag= c
FT /product= "Mature human secreted protein"
XX
XX WO200154708-A1.
PN
XX
XX 02-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US001434.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 18-AUG-2000; 2000US-0226279P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR;
PI Olsen HS, Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH;
PI Fiscella M, Ni J, Ruben SM, Barash SC;
XX
XX WPI; 2001-488743/53.
DR P-PSDB; AAE07060.
DR
XX
PT New isolated nucleic acids and polypeptides, useful for diagnosing,
PT treating and/or preventing human diseases and disorders.
XX
XX Claim 1; Page 452-453; 558pp; English.
PS
XX
CC AAD13345-AAD13401 represent cDNAs corresponding to 22 human secreted
CC protein genes, and AAE07051-AAE07105 represent the proteins they encode.
CC AAE07106-AAE07129 represent human secreted protein fragments or variants.
CC The genes and their secreted proteins are useful for preventing, treating
CC or ameliorating medical conditions, e.g., by protein or gene therapy.
CC Pathological conditions can be diagnosed by determining the amount of the
CC new protein in a sample or by determining the presence of mutations in
CC the new genes. Specific uses are described for each of the 22 genes,
CC based on the tissues in which they are most highly expressed, and include
CC developing products for the diagnosis or treatment of proliferative
CC disorders, cancer, tumours, foetal and developmental abnormalities,
CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,
CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
CC angiogenic disorders, kidney disorders, gastrointestinal disorders,
CC pregnancy-related disorders, endocrine disorders, and infections. The
CC proteins can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues,
CC to regenerate tissues, to identify their cognate ligands or binding
CC partners, and in chemotaxis, and can be used as a food additive or
CC preservative to modify storage properties. Antibodies specific for a
CC protein of the invention can be used in alleviating symptoms associated
CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
CC present sequence represents a human secreted protein-encoding cDNA of the
CC invention
XX
XX Sequence 1088 BP; 284 A; 246 C; 284 G; 269 T; 0 U; 5 Other;
SQ

Query Match

7.5%; Score 103; DB 4; Length 1088;

Best Local Similarity 100.0%; Pred. No. 1e-25;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGT 1321
Db |||||
QY 986 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGT 1045
Db |||||
QY 1322 ACGCGTGCATGCGACGTCATAGCTCTTCTATAGTGTACCTAA 1364
Db |||||
Db 1046 ACGCGTGCATGCGACGTCATAGCTCTTCTATAGTGTACCTAA 1088
RESULT 14
ABL90818/c
ID ABL90818 standard; cDNA; 2049 BP.
XX
AC ABL90818;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1380.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200190304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US016450.
XX
PR 19-MAY-2000; 2000US-0205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-122018/16.
DR P-PSDB; ABB90409.
XX
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders.
XX
PS Claim 4; SEQ ID NO 1380; 2081pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC (ABB9040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2049 BP; 383 A; 571 C; 686 G; 400 T; 0 U; 9 Other;
Query Match
7.2%; Score 99; DB 6; Length 2049;

Best Local Similarity 100.0%; Pred. No. 2e-24;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1271 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGTACGGGTGCA 1330
Db 321 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGTACGGGTGCA 262

QY 1331 TCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1369
Db 261 TCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 223

RESULT 15
ABZ55039
ID ABZ55039 standard; cDNA; 570 BP.
XX ABZ55039;
XX
DT 28-MAR-2003 (first entry)
XX
DE Aspergillus oryzae polynucleotide SEQ ID NO 4152.
XX
KW Aspergillus oryzae; fermentation; fungus; industrial; EST;
KW expressed sequence tag; gene; ss.
XX
OS Aspergillus oryzae.
XX
PN WO200279476-A1.
XX
PD 10-OCT-2002.
XX
PF 22-MAR-2002; 2002WO-IB000890.
XX
PR 30-MAR-2001; 2001JP-00098371.
XX
PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA (NARE-) NAT RES INST BREWING.
PA (NORQ) NAT FOOD RES INST MIN AGRIC.
XX
PI Machida M, Akita O, Kashiwagi Y, Kitamoto K, Horiuchi H;
PI Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;
XX
DR WPI; 2003-046817/04.

PT Detection of expression of specific Aspergillus genes for monitoring the
PT fermentation and growth conditions of the fungus, using DNA probes.
XX
PS Claim 1; SEQ ID NO 4152; 48pp + Sequence Listing; Japanese.
XX
CC The invention relates to a polynucleotide having any of 6006 specific
CC sequences (ABZ50888-ABZ56893), which are expressed by a fungus under
CC specific culture conditions including one or more of eutrophic,
CC oligotrophic, solid, early germination, alkaline, high temperature, low
CC temperature or maltose culture or polynucleotides stringently hybridising
CC to these sequences. The polynucleotides are useful for monitoring the
CC progress of fermentation and the growth conditions of a fungus,
CC especially of Aspergillus oryzae which is widely used in industrial
CC fermentation. Also monitoring for fungal contamination. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 570 BP; 144 A; 153 C; 147 G; 126 T; 0 U; 0 Other;

Query Match 7.1%; Score 98; DB 7; Length 570;
Best Local Similarity 100.0%; Pred. No. 5.8e-24;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1272 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGTACGGGTGCA 1331
Db 286 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGTACGGGTGCA 345

QY 1332 GCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1369

Db 346 GCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 383
Search completed: October 2, 2004, 06:11:50
Job time : 567 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2004, 05:28:56 ; Search time 5267 Seconds
(without alignments)
11290.420 Million cell updates/sec

Title: US-09-993-808B-1

Perfect score: 1372

Sequence: 1 cccacgcgtccggacgcaag.....agtgtaacctaattcattc 1372

Scoring table: OLIGO NUC

Gapop_60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID | Description |
|------------|-------|---------------|--------|----|-----------|--------------------|
| 1 | 108 | 7.9 | 2658 | 5 | DRE505990 | AJ505990 Danio rer |
| 2 | 98 | 7.1 | 2657 | 6 | BD275987 | BD275987 62 Human |
| 3 | 97 | 7.1 | 2454 | 6 | BD276004 | BD276004 62 Human |
| 4 | 95 | 6.9 | 1255 | 6 | AR137827 | AR137827 Sequence |
| 5 | 95 | 6.9 | 1255 | 6 | AR170186 | AR170186 Sequence |
| 6 | 95 | 6.9 | 1450 | 6 | BD247539 | BD247539 Methods f |
| 7 | 95 | 6.9 | 1581 | 6 | AR066494 | AR066494 Sequence |
| 8 | 95 | 6.9 | 1581 | 6 | I60018 | I60018 Sequence 1 |
| 9 | 95 | 6.9 | 2015 | 6 | BD136416 | BD136416 95 human |
| 10 | 95 | 6.9 | 2486 | 6 | AX301825 | AX301825 Sequence |
| 11 | 90 | 6.6 | 944 | 6 | AR184178 | AR184178 Sequence |
| 12 | 90 | 6.6 | 944 | 6 | BD192707 | BD192707 123 human |
| 13 | 89 | 6.5 | 1270 | 6 | BD275352 | BD275352 49 Human |
| 14 | 87 | 6.3 | 2581 | 6 | BD130531 | BD130531 Ligand re |
| 15 | 82 | 6.0 | 1330 | 6 | AR137825 | AR137825 Sequence |
| 16 | 82 | 6.0 | 1330 | 6 | AR170184 | AR170184 Sequence |
| 17 | 80 | 5.8 | 1297 | 8 | CSPATATLP | Y12793 Cucumis sat |
| 18 | 80 | 5.8 | 1984 | 6 | BD230989 | BD230989 49 human |
| 19 | 80 | 5.8 | 1984 | 6 | AR337284 | AR337284 Sequence |
| 20 | 79 | 5.8 | 204 | 6 | BD263433 | BD263433 Compositi |
| 21 | 79 | 5.8 | 204 | 6 | BD263434 | BD263434 Compositi |
| 22 | 79 | 5.8 | 270 | 20 | BTZ81343 | Z81343 B.taurus DN |
| 23 | 79 | 5.8 | 280 | 9 | HSA276253 | AJ276253 Homo sapi |
| 24 | 79 | 5.8 | 307 | 9 | HSA276255 | AJ276255 Homo sapi |
| 25 | 79 | 5.8 | 318 | 9 | HSA276256 | AJ276256 Homo sapi |
| 26 | 79 | 5.8 | 593 | 6 | AX780401 | AX780401 Sequence |
| 27 | 79 | 5.8 | 646 | 6 | AX067316 | AX067316 Sequence |
| 28 | 79 | 5.8 | 784 | 6 | AX780400 | AX780400 Sequence |
| 29 | 79 | 5.8 | 826 | 6 | AR339093 | AR339093 Sequence |
| 30 | 79 | 5.8 | 907 | 6 | AR339094 | AR339094 Sequence |
| 31 | 79 | 5.8 | 955 | 6 | AR338737 | AR338737 Sequence |
| 32 | 79 | 5.8 | 988 | 6 | AX644918 | AX644918 Sequence |
| 33 | 79 | 5.8 | 1188 | 6 | AX780374 | AX780374 Sequence |
| 34 | 79 | 5.8 | 1200 | 6 | AR339073 | AR339073 Sequence |
| 35 | 79 | 5.8 | 1495 | 6 | BD247537 | BD247537 Methods f |
| 36 | 79 | 5.8 | 2174 | 1 | AF250348 | AF250348 Carboxydo |
| 37 | 79 | 5.8 | 2322 | 6 | AX670965 | AX670965 Sequence |
| 38 | 79 | 5.8 | 2544 | 6 | AR175857 | AR175857 Sequence |
| 39 | 79 | 5.8 | 2644 | 6 | AX780580 | AX780580 Sequence |
| 40 | 79 | 5.8 | 2644 | 6 | AX780581 | AX780581 Sequence |
| 41 | 79 | 5.8 | 4109 | 12 | PSPORT1 | U12390 Cloning vec |
| 42 | 79 | 5.8 | 4431 | 12 | AY217101 | AY217101 Expressio |
| 43 | 79 | 5.8 | 5957 | 6 | BD263353 | BD263353 Compositi |
| 44 | 79 | 5.8 | 5957 | 6 | BD263354 | BD263354 Compositi |
| 45 | 79 | 5.8 | 6939 | 6 | AX780120 | AX780120 Sequence |

ALIGNMENTS

| | | | | | |
|------------|---|--|------|--------|-----------------|
| RESULT 1 | DRE505990 | 2658 bp | mrna | linear | VRT 12-AUG-2003 |
| LOCUS | DRE505990 | Danio rerio mRNA for heparan sulfate 6-O-sulfotransferase (hs6st | | | |
| DEFINITION | Danio rerio mRNA for heparan sulfate 6-O-sulfotransferase (hs6st | | | | |
| ACCESSION | AJ505990 | | | | |
| VERSION | AJ505990.1 | GI:29150000 | | | |
| KEYWORDS | heparan sulfate 6-O-sulfotransferase; hs6st gene. | | | | |
| SOURCE | Danio rerio (zebrafish) | | | | |
| ORGANISM | Danio rerio | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| AUTHORS | Actinopterygii; Neopterygii; Teleostei; Ostariophysi; | | | | |
| | Cypriniformes; Cyprinidae; Danio. | | | | |
| | 1 | | | | |
| | Bink,R.J., Habuhi,H., Lele,Z., Dolk,E., Joore,J., Rauch,G.J., | | | | |

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
CDS
ORIGIN
Query Match
Best Local Similarity
Matches
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
PN
PD
PF

Geisler,R., Wilson,S.W., de Hertog,J., Kimata,K. and Zivkovic,D.
Heparan Sulfate 6-O-Sulfotransferase Is Essential for Muscle
Development in Zebrafish
J. Biol. Chem. 278 (33), 31118-31127 (2003)
12782624
2 (bases 1 to 2658)
Bink,R.J.
Direct Submission
Submitted (19-AUG-2002) Bink R.J., Netherlands Inst. for Dev.
Biology, Hubrecht Laboratory, Uppsalaalan 8, 3584CT Utrecht,
NETHERLANDS
Location/Qualifiers
1. .2658
/organism="Danio rerio"
/mol_type="mrna"
/db_xref="taxon:7955"
/chromosome="14"
1. .2658
/gene="hs6st"
314. .1720
/gene="hs6st"
/function="brain and fin development, myogenesis"
/codon_start=1
/product="heparan sulfate 6-O-sulfotransferase"
/protein_id="CAD44529.1"
/db_xref="GI:29150001"
/db_xref="SPTREMBL:Q800H9"
/translation="MDGKSNYSRLLLALLMILFFGGIVLQYICSTSDWQLHLASLS
RLGSRAPGRLNGAGAGDPYSSDGLVRFVPRFNFTKDLRAVDPHIKGDDVIVFL
HIQKTGTTFGRLHVRNIQLERPCCHAGOKKCTYRPGKRDITWLSRFSTGWSGCLH
ADWTELTNCVPSFMSNRESQERRMTPSNYYITILRDPVWRYLSEWRHVQGTWKA
SKHMCDDGLPTLTLPSCYPGDDWSGSLFEFMVCPYNLANNROTMLADLSLVGCYN
LTVMSENQWAMLLLESAKRNLNMAFFGLTEYQRTQYLFTEHTRLSFIAPFTQLNGT
RAASVEVEPETQRRIRLELNQWDVELYEVARDLFLQRFQFARQQERREARQRIQERRK
LRKVKSWLGVTKAVFKPTKEPPMTQSPFAAEKQADAEQTLESETEGQVEENWLE
EDDGEIMLDYSENVEQWR"

7.9%; Score 108; DB 5; Length 2658;
100.0%; Pred. No. 3.8e-44;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGT 1321
|||||
2531 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGT 2590
|||||

1322 ACGCGTCATCGCAGCTCATAGCTTCTTATAGTGTCACTAAATTCA 1369
|||||
2591 ACGCGTCATCGCAGCTCATAGCTTCTTATAGTGTCACTAAATTCA 2638
|||||

BD275987 2657 bp DNA linear PAT 17-JUL-2003
62 Human secreted proteins.
BD275987
BD275987.1 GI:33085755
JP 2002543771-A/65.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2657)
Birse,C.E., Mouret,P.A., Florence,K.A., Ruben,S.M.,
Komatsoulis,G.A., Ni,J., Ebner,R., W.D., Lafleur, Olsen,H.S.,
Shi,Y., Soppet,D.R., Rosen,C.A. and Young,P.E.
62 Human secreted proteins
Patent: JP 2002543771-A 65 24-DEC-2002;
Human Genome Sciences Inc

Query Match
Best Local Similarity
Matches
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
PN
PD
PF

BD275987 2657 bp DNA linear PAT 17-JUL-2003
62 Human secreted proteins.
BD275987
BD275987.1 GI:33085755
JP 2002543771-A/65.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2657)
Birse,C.E., Mouret,P.A., Florence,K.A., Ruben,S.M.,
Komatsoulis,G.A., Ni,J., Ebner,R., W.D., Lafleur, Olsen,H.S.,
Shi,Y., Soppet,D.R., Rosen,C.A. and Young,P.E.
62 Human secreted proteins
Patent: JP 2002543771-A 65 24-DEC-2002;
Human Genome Sciences Inc

Query Match
Best Local Similarity
Matches
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
PN
PD
PF

BD275987 2657 bp DNA linear PAT 17-JUL-2003
62 Human secreted proteins.
BD275987
BD275987.1 GI:33085755
JP 2002543771-A/65.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2657)
Birse,C.E., Mouret,P.A., Florence,K.A., Ruben,S.M.,
Komatsoulis,G.A., Ni,J., Ebner,R., W.D., Lafleur, Olsen,H.S.,
Shi,Y., Soppet,D.R., Rosen,C.A. and Young,P.E.
62 Human secreted proteins
Patent: JP 2002543771-A 65 24-DEC-2002;
Human Genome Sciences Inc

PR 09-APR-1999 US 60/128693,26-APR-1999 US 60/130991 PI
Charles e birse,paul a mouret,kimberly a florence,steven m pi
ruben,
PI george a komatsoulis,jian ni,reinhard ebner,david w pi
lafleur,henrik s olsen,
PI yanggu shi,daniel r soppet,craig a rosen,paul e young cc
FH Key Location/Qualifiers
FT SITE (179)
FT /note='n equals a,t,g, or c'.
Location/Qualifiers
1. .2657
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
7.1%; Score 98; DB 6; Length 2657;
100.0%; Pred. No. 5.3e-39;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1272 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTCAT 1331
|||||
2307 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTCAT 2366
|||||

1332 GCGACGTCATAGCTTCTTATAGTGTCACTAAATTCA 1369
|||||
2367 GCGACGTCATAGCTTCTTATAGTGTCACTAAATTCA 2404
|||||

BD276004 2454 bp DNA linear PAT 17-JUL-2003
62 Human secreted proteins.
BD276004
BD276004.1 GI:33085772
JP 2002543771-A/82.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2454)
Birse,C.E., Mouret,P.A., Florence,K.A., Ruben,S.M.,
Komatsoulis,G.A., Ni,J., Ebner,R., W.D., Lafleur, Olsen,H.S.,
Shi,Y., Soppet,D.R., Rosen,C.A. and Young,P.E.
62 Human secreted proteins
Patent: JP 2002543771-A 82 24-DEC-2002;
Human Genome Sciences Inc

OS Homo sapiens
PN JP 2002543771-A/82
PD 24-DEC-2002
PF 06-APR-2000 JP 2000611564
PR 09-APR-1999 US 60/128693,26-APR-1999 US 60/130991 PI
Charles e birse,paul a mouret,kimberly a florence,steven m pi
ruben,
PI george a komatsoulis,jian ni,reinhard ebner,david w pi
lafleur,henrik s olsen,
PI yanggu shi,daniel r soppet,craig a rosen,paul e young cc
FH Key Location/Qualifiers
FT SITE (2317)
FT /note='n equals a,t,g, or c'.
Location/Qualifiers
1. .2454
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

1273 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTCATG 1332
|||||

Db 2105 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTGCA TG 2164

QY 1333 CGACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1369
|||||

Db 2165 CGACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 2201
|||||

RESULT 4
AR137827

LOCUS AR137827 1255 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 31 from patent US 6197561.
ACCESSION AR137827
VERSION AR137827.1 GI:14479336
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1255)
AUTHORS Martino-Catt,S.J., Wang,H., Beach,L.R., Wang,X. and Bowen,B.A.
TITLE Genes controlling phytate metabolism in plants and uses thereof
JOURNAL Patent: US 6197561-A 31 06-MAR-2001;
FEATURES
source Location/Qualifiers
1..1255
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 6.9%; Score 95; DB 6; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.9e-37;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1275 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTGCG 1334
|||||

Db 1161 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTGCG 1220
|||||

QY 1335 ACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1369
|||||

Db 1221 ACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1255
|||||

RESULT 5
AR170186

LOCUS AR170186 1255 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 31 from patent US 6291224.
ACCESSION AR170186
VERSION AR170186.1 GI:17908145
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1255)
AUTHORS Martino-Catt,S.J., Wang,H., Beach,L.R., Wang,X. and Bowen,B.A.
TITLE Genes controlling phytate metabolism in plants and uses thereof
JOURNAL Patent: US 6291224-A 31 18-SEP-2001;
FEATURES
source Location/Qualifiers
1..1255
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 6.9%; Score 95; DB 6; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.9e-37;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1275 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTGCG 1334
|||||

Db 1161 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTGCG 1220
|||||

QY 1335 ACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1369
|||||

Db 1221 ACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1255
|||||

RESULT 6
BD247539

LOCUS BD247539 1450 bp DNA linear PAT 17-JUL-2003
DEFINITION Methods for transforming plastids.
ACCESSION BD247539
VERSION BD247539.1 GI:33057309
KEYWORDS JP 2002531096-A/3.
SOURCE Brassica sp.
ORGANISM Brassica sp.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 1450)
AUTHORS Chaudhuri,S.
TITLE Methods for transforming plastids
JOURNAL Patent: JP 2002531096-A 3 24-SEP-2002;
COMMENT CALGENE LLC
OS Brassica sp. (mustard)
PN JP 2002531096-A/3
PD 24-SEP-2002
PE 24-NOV-1999 JP 2000585430
PR 25-NOV-1998 US 60/109892
PI SUMITA CHAUDHURI
PC C12N15/09,A01H5/00,C12N5/10,C12N15/00,C12N5/00 CC Methods
for transforming plastids
FH Key Location/Qualifiers
FT source 1..1450
FT Location/Qualifiers
1..1450
/organism="Brassica sp."
/mol_type="genomic DNA"
/db_xref="taxon:3717"

ORIGIN

Query Match 6.9%; Score 95; DB 6; Length 1450;
Best Local Similarity 100.0%; Pred. No. 1.9e-37;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1275 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTGCG 1334
|||||

Db 1346 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTGCG 1405
|||||

QY 1335 ACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1369
|||||

Db 1406 ACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1440
|||||

RESULT 7
AR066494

LOCUS AR066494 1581 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5850022.
ACCESSION AR066494
VERSION AR066494.1 GI:5996710
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1581)
AUTHORS Dehesh,K., Voelker,T. and Hawkins,D.
TITLE Production of myristate in plant cells
JOURNAL Patent: US 5850022-A 1 15-DEC-1998;
FEATURES
source Location/Qualifiers
1..1581
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 6.9%; Score 95; DB 6; Length 1581;
Best Local Similarity 100.0%; Pred. No. 1.9e-37;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1275 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGTACGCGTGCG 1334
|||||

Db 1479 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCAAAGCTTACGTACGGTGCATGCG 1538
1335 ACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1369
1539 ACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1573

RESULT 8
I60018 160018 1581 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 1 from patent US 5654495.
ACCESSION I60018
VERSION I60018.1 GI:2478650
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1581)
AUTHORS Voelker,T.Alois. and Davies,H.Maelor.
TITLE Production of myristate in plant cells
JOURNAL Patent: US 5654495-A 1 05-AUG-1997;
FEATURES Location/Qualifiers
source 1. .1581
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 6.9%; Score 95; DB 6; Length 1581;
Best Local Similarity 100.0%; Pred. No. 1.9e-37;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1275 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCAAAGCTTACGTACGGTGCATGCG 1334
Db 1479 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCAAAGCTTACGTACGGTGCATGCG 1538
QY 1335 ACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1369
Db 1539 ACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1573

RESULT 9
BD136416 2015 bp DNA linear PAT 18-SEP-2002
LOCUS BD136416 95 human secretory proteins.
DEFINITION BD136416
ACCESSION BD136416.1 GI:23231361
VERSION JP 2002506627-A/103.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2015)
AUTHORS Ruben,S.M., Ni,J., Rosen,C.A., Yu,G.L., Young,P.E., Fen,P.,
Soppet,D.R., Wei,Y.F., Endress,G.A., Duan,R.D., Kyaw,H., Ebner,R.,
Lafleur,D.W., Olsen,H.S., Shi,Y. and Moore,P.A.
95 human secretory proteins
Patent: JP 2002506627-A 103 05-MAR-2002;
HUMAN GENOME SCIENCES INC
OS Homo sapiens (human)
PN JP 2002506627-A/103
PD 05-MAR-2002
PF 18-MAR-1999 JP 2000536733
PR 19-MAR-1998 US 60/078566,19-MAR-1998 US 60/078576 PR
19-MAR-1998 US 60/078573,19-MAR-1998 US 60/078574 PR
19-MAR-1998 US 60/078579,19-MAR-1998 US 60/078578 PR
19-MAR-1998 US 60/078581,19-MAR-1998 US 60/078577 PR
19-MAR-1998 US 60/078563,01-APR-1998 US 60/080314 PR
01-APR-1998 US 60/080312,01-APR-1998 US 60/080313 PI
M RUBEN,JIAN NI, CRAIG A ROSEN, GUO
LIANG YU, PAUL E YOUNG,
PI PING FENG,
DANIEL R SOPPET, YING FEI WEI, GREGORY A ENDRESS, ROXANNE D DUAN,

PI HLA KYAW,
PI REINHARD EBNER, DAVID W LAFLEUR, HENRIK S OLSEN, YANGGU SHI, PAUL
PI A MOORE
PC C12N15/09, A61K38/00, A61K48/00, A61P43/00, C07K14/47, C07K16/18,
PC C12N1/15,
PC C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12Q1/68, G01N33/53, G01N33/ PC
566,
PC C12N15/00, A61K37/02, C12N5/00
CC n equals a,t,g, or c
CC n equals a,t,g, or c
CC n equals a,t,g, or c
CC n equals a,t,g, or c
CC n equals a,t,g, or c
CC n equals a,t,g, or c
CC n equals a,t,g, or c
FH Key Location/Qualifiers
FT source 1. .2015
FT /organism="Homo sapiens (human)"
FT Location/Qualifiers
1. .2015
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 6.9%; Score 95; DB 6; Length 2015;
Best Local Similarity 100.0%; Pred. No. 1.9e-37;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1262 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCAAAGCTTACGT 1321
Db 1548 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGATCCAAAGCTTACGT 1607
QY 1322 ACGCGTGCATGCGACGTACATAGCTCTTCTATAGTG 1356
Db 1608 ACGCGTGCATGCGACGTACATAGCTCTTCTATAGTG 1642

RESULT 10
AX301825 2486 bp DNA linear PAT 30-NOV-2001
LOCUS AX301825
DEFINITION Sequence 4 from Patent WO0185768.
ACCESSION AX301825
VERSION AX301825.1 GI:17382883
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Gu,W.
TITLE Novel g-protein coupled receptors and uses therefor
JOURNAL Patent: WO 0185768-A 4 15-NOV-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES Location/Qualifiers
source 1. .2486
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS <2. .1903
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD13077.1"
/db_xref="GI:17382884"
/translation="NTHYRESWYACRYSGIPGSPSTHASVERSOGLSLPAHPASLAAL
AASNTTASGKLELTPSQLSSQLALDLSWNAIRSIHPEAFSTLHSLVKLDLTDNQLTT
LPLAGLGLMLHLKGNLALSQAFSKDSFPLRLILEVPYAYQCCPYGMCASFFKASGQ
WEAEDLHLDDESSKPLGLLARQENHYDQDLDELQLEMEDSKPHPSVCSPTPGF
KPCEYLPESWGIRLAVVAIVLLSVLCNGLVLLTVFAGGPAPLPVVKFVVGAIAAGNTL
TGISCGLLASVDALTFQFSEYGARWETGLGCRATGLAVLGSSEASVLLTLTAAVQCS
VSVSCVRAYGKSPSLGSRAGVLGCLALAGLAAALPLASVGEYGASPLCLPYAPPEGQ
PAALGFTVALVMNSFCLVVGAYIKLYCDLPRGDFEAWDCAMVRHVAVLIFADGL

FT SITE (553)
/note='n equals a,t,g, or c'
FT SITE (1070)
/note='n equals a,t,g, or c'
FT SITE (1208)
/note='n equals a,t,g, or c'.
FEATURES Location/Qualifiers
source 1. .1270
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 6.5%; Score 89; DB 6; Length 1270;
Best Local Similarity 100.0%; Pred. No. 2.3e-34;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1270 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGCGGTGC 1329
|||||
Db 981 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGCGGTGC 1040
QY 1330 ATGCGACGTCATAGCTCTTCTATAGTGTC 1358
|||||
Db 1041 ATGCGACGTCATAGCTCTTCTATAGTGTC 1069

RESULT 14
BD130531 2581 bp DNA linear PAT 18-SEP-2002
LOCUS
DEFINITION Ligand receptors and utilization thereof.
ACCESSION BD130531
VERSION BD130531.1 GI:23225476
KEYWORDS JP 2002501083-A/1.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2581)
AUTHORS Glucksmann,A.M. and Robison,K.
TITLE Ligand receptors and utilization thereof
JOURNAL Patent: JP 2002501083-A 1 15-JAN-2002;
MILLENNIUM PHARMACEUTICALS INC
COMMENT OS Unidentified
PN JP 2002501083-A/1
PD 15-JAN-2002
PF 20-JAN-1999 JP 2000528599
PR 26-JAN-1998 US 09/013634
PI ALEXANDRA M GLUCKSMANN,KEITH ROBISON
PC C07K14/705,C07K16/28,C12N5/10,C12N15/09,C12P21/02,C12Q1/68, PC
G01N33/53,
PC C12N5/00,C12N15/00
CC Strandedness: Single;
CC Topology: Linear;
CC Ligand receptors and utilization thereof
FH Key Location/Qualifiers
FT CDS 184..1194.

FEATURES Location/Qualifiers
source 1. .2581
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 6.3%; Score 87; DB 6; Length 2581;
Best Local Similarity 100.0%; Pred. No. 2.4e-33;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1262 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGT 1321
|||||
Db 2492 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGT 2551
QY 1322 ACGCGTGATCGACGTCATAGCTTT 1348
|||||
Db 2552 ACGCGTGATCGACGTCATAGCTTT 2578

RESULT 15
AR137825/c 1330 bp DNA linear PAT 16-JUN-2001
LOCUS
DEFINITION Sequence 29 from patent US 6197561.
ACCESSION AR137825
VERSION AR137825.1 GI:14479334
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1330)
AUTHORS Martino-Catt,S.J., Wang,H., Beach,L.R., Wang,X. and Bowen,B.A.
TITLE Genes controlling phytate metabolism in plants and uses thereof
JOURNAL Patent: US 6197561-A 29 06-MAR-2001;
FEATURES Location/Qualifiers
source 1. .1330
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 6.0%; Score 82; DB 6; Length 1330;
Best Local Similarity 100.0%; Pred. No. 9.3e-31;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1270 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGCGGTGC 1329
|||||
Db 94 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAGCTTACGCGGTGC 35
QY 1330 ATGCGACGTCATAGCTCTTCTA 1351
|||||
Db 34 ATGCGACGTCATAGCTCTTCTA 13

Search completed: October 2, 2004, 07:39:50
Job time : 5271 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2004, 05:49:02 ; Search time 127 Seconds
(without alignments)
5995.218 Million cell updates/sec

Title: US-09-993-808B-1
Perfect score: 1372
Sequence: 1 cccacgcgtccggacgaag.....agtgcaacctaattcattc 1372

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0
Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 95 | 6.9 | 1255 | 3 | US-09-118-442-31 |
| 2 | 95 | 6.9 | 1255 | 3 | US-09-677-064-31 |
| 3 | 95 | 6.9 | 1581 | 1 | US-08-383-756-1 |
| 4 | 95 | 6.9 | 1581 | 2 | US-08-460-898-1 |
| 5 | 90 | 6.6 | 944 | 4 | US-09-227-357-122 |
| 6 | 87 | 6.3 | 2581 | 2 | US-09-013-634-1 |
| 7 | 82 | 6.0 | 1330 | 3 | US-09-118-442-29 |
| 8 | 82 | 6.0 | 1330 | 3 | US-09-677-064-29 |
| 9 | 80 | 5.8 | 1984 | 4 | US-09-904-615-46 |
| 10 | 79 | 5.8 | 826 | 4 | US-09-620-312D-584 |
| 11 | 79 | 5.8 | 907 | 4 | US-09-620-312D-585 |
| 12 | 79 | 5.8 | 955 | 4 | US-09-620-312D-228 |
| 13 | 79 | 5.8 | 1200 | 4 | US-09-620-312D-564 |
| 14 | 79 | 5.8 | 2544 | 4 | US-09-483-371-1 |
| 15 | 79 | 5.8 | 2544 | 4 | US-09-957-156-1 |
| 16 | 78 | 5.7 | 1243 | 4 | US-09-620-312D-1092 |
| 17 | 78 | 5.7 | 2638 | 3 | US-09-042-785A-22 |
| 18 | 77 | 5.6 | 1689 | 4 | US-09-053-374A-4 |
| 19 | 76 | 5.5 | 657 | 4 | US-09-620-312D-543 |
| 20 | 76 | 5.5 | 1067 | 4 | US-09-620-312D-853 |
| 21 | 76 | 5.5 | 1663 | 4 | US-09-620-312D-6 |
| 22 | 76 | 5.5 | 2085 | 4 | US-09-620-312D-1002 |
| 23 | 76 | 5.5 | 3396 | 4 | US-09-668-680-6 |
| 24 | 76 | 5.5 | 3423 | 4 | US-09-668-680-7 |
| 25 | 75 | 5.5 | 831 | 4 | US-09-904-615-25 |
| 26 | 75 | 5.5 | 2202 | 4 | US-09-396-149-3 |
| 27 | 73 | 5.3 | 322 | 4 | US-08-956-171E-1520 |

| | | | | | | |
|----|----|-----|------|---|--------------------|-------------------|
| 28 | 73 | 5.3 | 533 | 4 | US-09-105-542A-11 | Sequence 11, Appl |
| 29 | 73 | 5.3 | 2727 | 4 | US-09-620-312D-218 | Sequence 218, App |
| 30 | 73 | 5.3 | 3327 | 3 | US-08-689-421-26 | Sequence 26, Appl |
| 31 | 73 | 5.3 | 3327 | 3 | US-09-389-528-26 | Sequence 26, Appl |
| 32 | 73 | 5.3 | 3327 | 3 | US-09-181-827A-26 | Sequence 26, Appl |
| 33 | 72 | 5.2 | 1790 | 4 | US-09-118-637A-3 | Sequence 3, Appli |
| 34 | 71 | 5.2 | 2460 | 3 | US-08-964-127-1 | Sequence 1, Appli |
| 35 | 71 | 5.2 | 2460 | 4 | US-09-496-692-1 | Sequence 1, Appli |
| 36 | 71 | 5.2 | 2460 | 4 | US-10-000-273-1 | Sequence 1, Appli |
| 37 | 69 | 5.0 | 1746 | 4 | US-09-489-847-21 | Sequence 21, Appl |
| 38 | 68 | 5.0 | 1079 | 4 | US-09-482-273-61 | Sequence 61, Appl |
| 39 | 68 | 5.0 | 1111 | 4 | US-09-620-312D-797 | Sequence 797, App |
| 40 | 68 | 5.0 | 2311 | 4 | US-09-489-847-123 | Sequence 123, App |
| 41 | 67 | 4.9 | 774 | 4 | US-09-105-542A-4 | Sequence 4, Appli |
| 42 | 67 | 4.9 | 2664 | 4 | US-09-149-476-255 | Sequence 255, App |
| 43 | 65 | 4.7 | 1006 | 3 | US-08-911-423-3 | Sequence 3, Appli |
| 44 | 64 | 4.7 | 501 | 4 | US-09-439-313-451 | Sequence 451, App |
| 45 | 64 | 4.7 | 501 | 4 | US-09-352-616A-451 | Sequence 451, App |

ALIGNMENTS

RESULT 1
US-09-118-442-31
; Sequence 31, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; APPLICANT: Bowen, Benjamin A.
; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; CURRENT FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; EARLIER FILING DATE: 1997-07-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 1255
; TYPE: DNA
; ORGANISM: Zea mays
US-09-118-442-31

| | | | | |
|-----------------------|--------|--|----------|---------------------------------|
| Query Match | 6.9% | Score 95; | DB 3; | Length 1255; |
| Best Local Similarity | 100.0% | Pred. No. | 7.3e-28; | |
| Matches | 95; | Conservative | 0; | Mismatches 0; Indels 0; Gaps 0; |
| QY | 1275 | AAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGGTGCATGCG | 1334 | |
| Db | 1161 | AAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGGTGCATGCG | 1220 | |
| QY | 1335 | ACGTACATAGCTTCTTCTATAGTGTCACTAAATTCA | 1369 | |
| Db | 1221 | ACGTACATAGCTTCTTCTATAGTGTCACTAAATTCA | 1255 | |

RESULT 2
US-09-677-064-31
; Sequence 31, Application US/09677064
; Patent No. 6291224
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.

;; TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
;; TITLE OF INVENTION: Plants and Uses Thereof

;; FILE REFERENCE: 0706D
;; CURRENT APPLICATION NUMBER: US/09/677,064
;; CURRENT FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/055,446
;; PRIOR FILING DATE: 1997-08-11
;; PRIOR APPLICATION NUMBER: 60/055,526
;; PRIOR FILING DATE: 1997-08-08
;; PRIOR APPLICATION NUMBER: 60/053,944
;; PRIOR FILING DATE: 1997-07-28
;; PRIOR APPLICATION NUMBER: 09/118,442
;; PRIOR FILING DATE: 1998-07-17

;; NUMBER OF SEQ ID NOS: 31
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 31
;; LENGTH: 1255
;; TYPE: DNA
;; ORGANISM: Zea mays
US-09-677-064-31

Query Match 6.9%; Score 95; DB 3; Length 1255;
Best Local Similarity 100.0%; Pred. No. 7.3e-28;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1275 AAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGCATGCG 1334
Db 1161 AAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGCATGCG 1220
QY 1335 ACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1369
Db 1221 ACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1255

RESULT 3

US-08-383-756-1
; Sequence 1, Application US/08383756
; Patent No. 5654495

GENERAL INFORMATION:

; APPLICANT: Dehesh, Katayoon
; APPLICANT: Voelker, Toni Alois
; APPLICANT: Hawkins, Deborah
; APPLICANT: Davies, Huw Maelor
; TITLE OF INVENTION: Production of Myristate in Plant Cells
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 5.1(a)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/383,756
; FILING DATE: 02-FEB-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/261,695
; FILING DATE: 16-JUN-94
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10814
; FILING DATE: 29-OCT-93
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 07/968,971
; FILING DATE: 30-OCT-92
; ATTORNEY/AGENT INFORMATION:

; NAME: Elizabeth Lassen
; REGISTRATION NUMBER: 31,845
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1581 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
US-08-383-756-1

Query Match 6.9%; Score 95; DB 1; Length 1581;
Best Local Similarity 100.0%; Pred. No. 7e-28;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1275 AAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGCATGCG 1334
Db 1479 AAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGCATGCG 1538
QY 1335 ACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1369
Db 1539 ACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1573

RESULT 4

US-08-460-898-1
; Sequence 1, Application US/08460898
; Patent No. 5850022

GENERAL INFORMATION:

; APPLICANT: Dehesh, Katayoon
; APPLICANT: Voelker, Toni Alois
; APPLICANT: Hawkins, Deborah
; APPLICANT: Davies, Huw Maelor
; TITLE OF INVENTION: Production of Myristate in Plant Cells
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 5.1(a)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,898
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/383,756
; FILING DATE: 02-FEB-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/261,695
; FILING DATE: 16-JUN-94
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10814
; FILING DATE: 29-OCT-93
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 07/968,971

```
/ FILING DATE: 30-OCT-92
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Elizabeth Lassen
/ REGISTRATION NUMBER: 31,845
/ NAME: Donna E. Scherer
/ REGISTRATION NUMBER: 34,719
/ NAME: Carl J. Schwedler
/ REGISTRATION NUMBER: 36,924
/ REFERENCE/DOCKET NUMBER: CGNE 111-1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (916) 753-6313
/ TELEFAX: (916) 753-1510
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1581 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA to mRNA
US-08-460-898-1

Query Match          6.9%; Score 95; DB 2; Length 1581;
Best Local Similarity 100.0%; Pred. No. 7e-28;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1275 AAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGTACGCGTGATGCG 1334
Db 1479 AAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGTACGCGTGATGCG 1538

QY 1335 ACGTCATAGCTCTTCTATAGTGTCACTAAATTCA 1369
Db 1539 ACGTCATAGCTCTTCTATAGTGTCACTAAATTCA 1573
```

RESULT 5

```
US-09-227-357-122
; Sequence 122, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: PZ010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; EARLIER FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
```

```
/ EARLIER FILING DATE: 1997-07-08
/ EARLIER APPLICATION NUMBER: 60/051,919
/ EARLIER FILING DATE: 1997-07-08
/ EARLIER APPLICATION NUMBER: 60/051,928
/ EARLIER FILING DATE: 1997-07-08
/ EARLIER APPLICATION NUMBER: 60/055,722
/ EARLIER FILING DATE: 1997-08-18
/ EARLIER APPLICATION NUMBER: 60/055,723
/ EARLIER FILING DATE: 1997-08-18
/ EARLIER APPLICATION NUMBER: 60/055,948
/ EARLIER FILING DATE: 1997-08-18
/ EARLIER APPLICATION NUMBER: 60/055,949
/ EARLIER FILING DATE: 1997-08-18
/ EARLIER APPLICATION NUMBER: 60/055,953
/ EARLIER FILING DATE: 1997-08-18
/ EARLIER APPLICATION NUMBER: 60/055,950
/ EARLIER FILING DATE: 1997-08-18
/ EARLIER APPLICATION NUMBER: 60/055,947
/ EARLIER FILING DATE: 1997-08-18
/ EARLIER APPLICATION NUMBER: 60/055,964
/ EARLIER FILING DATE: 1997-08-18
/ EARLIER APPLICATION NUMBER: 60/056,360
/ EARLIER FILING DATE: 1997-08-18
/ EARLIER APPLICATION NUMBER: 60/055,684
/ EARLIER FILING DATE: 1997-08-18
/ EARLIER APPLICATION NUMBER: 60/055,984
/ EARLIER FILING DATE: 1997-08-18
/ EARLIER APPLICATION NUMBER: 60/055,954
/ EARLIER FILING DATE: 1997-08-18
/ EARLIER APPLICATION NUMBER: 60/058,785
/ EARLIER FILING DATE: 1997-09-12
/ EARLIER APPLICATION NUMBER: 60/058,664
/ EARLIER FILING DATE: 1997-09-12
/ EARLIER APPLICATION NUMBER: 60/058,660
/ EARLIER FILING DATE: 1997-09-12
/ EARLIER APPLICATION NUMBER: 60/058,661
/ EARLIER FILING DATE: 1997-09-12
/ NUMBER OF SEQ ID NOS: 672
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 122
/ LENGTH: 944
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (932)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (942)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (944)
/ OTHER INFORMATION: n equals a,t,g, or c
US-09-227-357-122

Query Match          6.6%; Score 90; DB 4; Length 944;
Best Local Similarity 100.0%; Pred. No. 6.5e-26;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGT 1321
Db 835 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAGCTTACGT 894

QY 1322 ACGCGTGCATCGGACGTCATAGCTCTTCTA 1351
Db 895 ACGCGTGCATCGGACGTCATAGCTCTTCTA 924

RESULT 6
US-09-013-634-1
; Sequence 1, Application US/09013634
```


Patent No. 5945307
GENERAL INFORMATION:
APPLICANT: M. Alexandra Glucksmann and Keith Robison
TITLE OF INVENTION: LIGAND RECEPTORS AND USES THEREFOR
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/013,634
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Attorney, Jean M. Silveri
REGISTRATION NUMBER: 39,030
REFERENCE/DOCKET NUMBER: MNI-036
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2581 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 184..1194
US-09-013-634-1

Query Match 6.3%; Score 87; DB 2; Length 2581;
Best Local Similarity 100.0%; Pred. No. 7.7e-25;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAGCTTACGT 1321
| | | | |
Db 2492 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAGCTTACGT 2551
| | | | |
QY 1322 ACGCGTGCATGCGACGTCATAGCTCTT 1348
| | | | |
Db 2552 ACGCGTGCATGCGACGTCATAGCTCTT 2578
| | | | |

RESULT 7
US-09-118-442-29/c
Sequence 29, Application US/09118442B
Patent No. 6197561
GENERAL INFORMATION:
APPLICANT: Martino-Catt, Susan J.
APPLICANT: Wang, Hongyu
APPLICANT: Beach, Larry R.
APPLICANT: Wang, Xun
APPLICANT: Bowen, Benjamin A.
TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
TITLE OF INVENTION: Plants and Uses Thereof
FILE REFERENCE: 0706
CURRENT APPLICATION NUMBER: US/09/118,442B
CURRENT FILING DATE: 1998-07-17
EARLIER APPLICATION NUMBER: 60/055,446
EARLIER FILING DATE: 1997-08-11

EARLIER APPLICATION NUMBER: 60/055,526
EARLIER FILING DATE: 1997-08-08
EARLIER APPLICATION NUMBER: 60/053,944
EARLIER FILING DATE: 1997-07-28
NUMBER OF SEQ ID NOS: 31
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 29
LENGTH: 1330
TYPE: DNA
ORGANISM: Zea mays
US-09-118-442-29

Query Match 6.0%; Score 82; DB 3; Length 1330;
Best Local Similarity 100.0%; Pred. No. 7.3e-23;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1270 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAGCTTACGTACGCGTGC 1329
| | | | |
Db 94 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAGCTTACGTACGCGTGC 35
| | | | |
QY 1330 ATGCGACGTCATAGCTTCTTA 1351
| | | | |
Db 34 ATGCGACGTCATAGCTTCTTA 13

RESULT 8
US-09-677-064-29/c
Sequence 29, Application US/09677064
Patent No. 6291224
GENERAL INFORMATION:
APPLICANT: Martino-Catt, Susan J.
APPLICANT: Wang, Hongyu
APPLICANT: Beach, Larry R.
TITLE OF INVENTION: Genes Controlling Phytate Metabolism in
TITLE OF INVENTION: Plants and Uses Thereof
FILE REFERENCE: 0706D
CURRENT APPLICATION NUMBER: US/09/677,064
CURRENT FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/055,446
PRIOR FILING DATE: 1997-08-11
PRIOR APPLICATION NUMBER: 60/055,526
PRIOR FILING DATE: 1997-08-08
PRIOR APPLICATION NUMBER: 60/053,944
PRIOR FILING DATE: 1997-07-28
PRIOR APPLICATION NUMBER: 09/118,442
PRIOR FILING DATE: 1998-07-17
NUMBER OF SEQ ID NOS: 31
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 29
LENGTH: 1330
TYPE: DNA
ORGANISM: Zea mays
US-09-677-064-29

Query Match 6.0%; Score 82; DB 3; Length 1330;
Best Local Similarity 100.0%; Pred. No. 7.3e-23;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1270 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAGCTTACGTACGCGTGC 1329
| | | | |
Db 94 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAGCTTACGTACGCGTGC 35
| | | | |
QY 1330 ATGCGACGTCATAGCTTCTTA 1351
| | | | |
Db 34 ATGCGACGTCATAGCTTCTTA 13

RESULT 9
US-09-904-615-46
Sequence 46, Application US/09904615
Patent No. 6566325
GENERAL INFORMATION:
APPLICANT: Rosen et al.

```

; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: PZ032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; CURRENT FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
; LENGTH: 1984
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (106)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-904-615-46

Query Match
Best Local Similarity 5.8%; Score 80; DB 4; Length 1984;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1275 AAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGCATGCG 1334
Db 1898 AAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGCATGCG 1957

QY 1335 ACGTCATAGCTCTTCTATAG 1354
Db 1958 ACGTCATAGCTCTTCTATAG 1977

RESULT 10
US-09-620-312D-584
; Sequence 584, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 584
; LENGTH: 907
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (225) .. (695)
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 584
; LENGTH: 826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; NAME/KEY: CDS
; LOCATION: (225) .. (614)
US-09-620-312D-584

Query Match
Best Local Similarity 5.8%; Score 79; DB 4; Length 826;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1291 GCGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGCATGCGACGTCTCTTCT 1350
Db 743 GCGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGCATGCGACGTCTCTTCT 802

QY 1351 ATAGTGTACCTAAATTCA 1369
Db 803 ATAGTGTACCTAAATTCA 821

RESULT 11
US-09-620-312D-585
; Sequence 585, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 585
; LENGTH: 907
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (225) .. (695)
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 585
; LENGTH: 907
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (225) .. (695)
US-09-620-312D-585

Query Match
Best Local Similarity 5.8%; Score 79; DB 4; Length 907;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1291 GCGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGCATGCGACGTCTCTTCT 1350
Db 824 GCGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGCATGCGACGTCTCTTCT 883

QY 1351 ATAGTGTACCTAAATTCA 1369
Db 884 ATAGTGTACCTAAATTCA 902

RESULT 12

```

```
US-09-620-312D-228/c
; Sequence 228, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yuning
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 228
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (80)..(583)
US-09-620-312D-228

Query Match      5.8%; Score 79; DB 4; Length 955;
Best Local Similarity 100.0%; Pred. No. 1.1e-21;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1291 GCGGGCGGCTCTAGAGGATCCAAAGCTTACGTACGGGTGCATGCGACGTCTCTTCT 1350
Db 120 GCGGGCGGCTCTAGAGGATCCAAAGCTTACGTACGGGTGCATGCGACGTCTCTTCT 61

QY 1351 ATAGTGTACCTAAATTCA 1369
Db 60 ATAGTGTACCTAAATTCA 42

RESULT 13
US-09-620-312D-564/c
; Sequence 564, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yuning
; APPLICANT: Wang, Dunrui
```

```
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 564
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (559)..(957)
US-09-620-312D-564

Query Match      5.8%; Score 79; DB 4; Length 1200;
Best Local Similarity 100.0%; Pred. No. 1.1e-21;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1291 GCGGGCGGCTCTAGAGGATCCAAAGCTTACGTACGGGTGCATGCGACGTCTCTTCT 1350
Db 86 GCGGGCGGCTCTAGAGGATCCAAAGCTTACGTACGGGTGCATGCGACGTCTCTTCT 27

QY 1351 ATAGTGTACCTAAATTCA 1369
Db 26 ATAGTGTACCTAAATTCA 8

RESULT 14
US-09-483-371-1
; Sequence 1, Application US/09483371
; Patent No. 6309869
; GENERAL INFORMATION:
; APPLICANT: Debbie S. Yaver
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; TITLE OF INVENTION: Polypeptides Having Acid Phosphatase
; TITLE OF INVENTION: Activity And Nucleic Acids Encoding Same
; FILE REFERENCE: 5791.200-US
; CURRENT APPLICATION NUMBER: US/09/483,371
; CURRENT FILING DATE: 2000-01-14
; EARLIER APPLICATION NUMBER: 09/231,612
; EARLIER FILING DATE: 1999-01-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2544
; TYPE: DNA
; ORGANISM: Fusarium
US-09-483-371-1

Query Match      5.8%; Score 79; DB 4; Length 2544;
Best Local Similarity 100.0%; Pred. No. 9.2e-22;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1291 GCGGGCGGCTCTAGAGGATCCAAAGCTTACGTACGGGTGCATGCGACGTCTCTTCT 1350
Db 2305 GCGGGCGGCTCTAGAGGATCCAAAGCTTACGTACGGGTGCATGCGACGTCTCTTCT 2364

QY 1351 ATAGTGTACCTAAATTCA 1369
Db 2365 ATAGTGTACCTAAATTCA 2383

RESULT 15
```

US-09-957-156-1
; Sequence 1, Application US/09957156
; Patent No. 6667169
; GENERAL INFORMATION:
; APPLICANT: Debbie S. Yaver
; APPLICANT: Randy M. Berka
; APPLICANT: Michael W. Rey
; TITLE OF INVENTION: Polypeptides Having Acid Phosphatase
; TITLE OF INVENTION: Activity And Nucleic Acids Encoding Same
; FILE REFERENCE: 5791.200-US
; CURRENT APPLICATION NUMBER: US/09/957,156
; CURRENT FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/483,371
; PRIOR FILING DATE: EARLIER FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2544
; TYPE: DNA
; ORGANISM: Fusarium
US-09-957-156-1

Query Match 5.8%; Score 79; DB 4; Length 2544;
Best Local Similarity 100.0%; Pred. No. 9.2e-22;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1291 GGCGGCCGCTCTAGAGGATCCAAGCTTACGTACGCGTGCATGCGACGTCATAGCTTCT 1350
Db |||||
Db 2305 GGCGGCCGCTCTAGAGGATCCAAGCTTACGTACGCGTGCATGCGACGTCATAGCTTCT 2364
QY 1351 ATAGTGTCACCTAAATTCA 1369
Db |||||
Db 2365 ATAGTGTCACCTAAATTCA 2383

Search completed: October 2, 2004, 08:37:50
Job time : 128 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2004, 07:39:55 ; Search time 702 Seconds
(without alignments)
9908.042 Million cell updates/sec

Title: US-09-993-808B-1
Perfect score: 1372
Sequence: 1 cccacgcgtccggacgcaag.....agtgtaacctaattcattc 1372

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3340653 seqs, 2534783454 residues

Word size : 0
Total number of hits satisfying chosen parameters: 6681306

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|----|---------------------|
| 1 | 1372 | 100.0 | 1372 | 9 | US-09-993-308-1 |
| 2 | 1372 | 100.0 | 1372 | 11 | US-09-993-808B-1 |
| 3 | 360 | 26.2 | 985 | 13 | US-10-425-114-18379 |
| 4 | 360 | 26.2 | 1111 | 13 | US-10-425-114-34947 |
| 5 | 108 | 7.9 | 149 | 10 | US-09-764-891-2570 |
| 6 | 108 | 7.9 | 1557 | 9 | US-09-925-297-262 |
| 7 | 108 | 7.9 | 1557 | 15 | US-10-023-896-42 |
| 8 | 108 | 7.9 | 2116 | 9 | US-09-925-297-261 |
| 9 | 108 | 7.9 | 2116 | 15 | US-10-023-896-31 |
| 10 | 106 | 7.7 | 701 | 15 | US-10-106-698-519 |
| 11 | 103 | 7.5 | 845 | 16 | US-10-264-049-133 |
| 12 | 99 | 7.2 | 2049 | 16 | US-10-264-237-1380 |
| 13 | 98 | 7.1 | 2657 | 15 | US-10-050-704-66 |
| 14 | 98 | 7.1 | 2657 | 17 | US-10-798-512-66 |

| | | | | | | |
|----|----|-----|------|----|--------------------|-------------------|
| 15 | 97 | 7.1 | 544 | 15 | US-10-106-698-2395 | Sequence 2395, Ap |
| 16 | 97 | 7.1 | 652 | 16 | US-10-264-049-1845 | Sequence 1845, Ap |
| 17 | 97 | 7.1 | 2454 | 15 | US-10-050-704-93 | Sequence 93, Appl |
| 18 | 97 | 7.1 | 2454 | 17 | US-10-798-512-93 | Sequence 93, Appl |
| 19 | 96 | 7.0 | 377 | 15 | US-10-080-254-14 | Sequence 14, Appl |
| 20 | 96 | 7.0 | 377 | 16 | US-10-242-355-17 | Sequence 17, Appl |
| 21 | 96 | 7.0 | 452 | 16 | US-10-264-049-1603 | Sequence 1603, Ap |
| 22 | 96 | 7.0 | 762 | 16 | US-10-264-049-1803 | Sequence 1803, Ap |
| 23 | 96 | 7.0 | 894 | 10 | US-09-969-730-17 | Sequence 17, Appl |
| 24 | 96 | 7.0 | 894 | 17 | US-10-621-363-17 | Sequence 17, Appl |
| 25 | 96 | 7.0 | 1138 | 9 | US-09-925-300-702 | Sequence 702, App |
| 26 | 96 | 7.0 | 1958 | 15 | US-10-106-698-2092 | Sequence 2092, Ap |
| 27 | 96 | 7.0 | 1958 | 16 | US-10-264-049-809 | Sequence 809, App |
| 28 | 96 | 7.0 | 1958 | 16 | US-10-264-237-1264 | Sequence 1264, Ap |
| 29 | 95 | 6.9 | 395 | 15 | US-10-106-698-2389 | Sequence 2389, Ap |
| 30 | 95 | 6.9 | 498 | 9 | US-09-925-299-331 | Sequence 331, App |
| 31 | 95 | 6.9 | 498 | 10 | US-09-925-299-331 | Sequence 331, App |
| 32 | 95 | 6.9 | 606 | 16 | US-10-264-049-1806 | Sequence 1806, Ap |
| 33 | 95 | 6.9 | 611 | 10 | US-09-764-872-108 | Sequence 108, App |
| 34 | 95 | 6.9 | 657 | 16 | US-10-264-049-1780 | Sequence 1780, Ap |
| 35 | 95 | 6.9 | 661 | 16 | US-10-264-049-1804 | Sequence 1804, Ap |
| 36 | 95 | 6.9 | 670 | 16 | US-10-264-049-465 | Sequence 465, App |
| 37 | 95 | 6.9 | 684 | 9 | US-09-925-299-320 | Sequence 320, App |
| 38 | 95 | 6.9 | 684 | 10 | US-09-925-299-320 | Sequence 320, App |
| 39 | 95 | 6.9 | 702 | 9 | US-09-925-301-732 | Sequence 732, App |
| 40 | 95 | 6.9 | 719 | 16 | US-10-264-049-1776 | Sequence 1776, Ap |
| 41 | 95 | 6.9 | 762 | 16 | US-10-264-049-1802 | Sequence 1802, Ap |
| 42 | 95 | 6.9 | 821 | 16 | US-10-264-049-1805 | Sequence 1805, Ap |
| 43 | 95 | 6.9 | 1188 | 15 | US-10-106-698-1637 | Sequence 1637, Ap |
| 44 | 95 | 6.9 | 1195 | 16 | US-10-158-057-90 | Sequence 90, Appl |
| 45 | 95 | 6.9 | 1255 | 9 | US-09-921-232-31 | Sequence 31, Appl |

ALIGNMENTS

RESULT 1
US-09-993-308-1
; Sequence 1, Application US/099933308
; Patent No. US20020159435A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William J.
; APPLICANT: Lowe, Keith S.
; APPLICANT: Larkins, Brian A.
; APPLICANT: Dilkes, Brian R.
; APPLICANT: Sun, Yuejin
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,308
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 60/246,349
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1372
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (134)..(904)
; OTHER INFORMATION:
US-09-993-308-1

Query Match 100.0%; Score 1372; DB 9; Length 1372;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCCACGGTCCGACGCAAGCGGCTGCAGCGCAGCGCGCGCGCGCTGTGGCCTGT 60
Db 1 CCCACGGTCCGACGCAAGCGGCTGCAGCGCAGCGCGCGCGCGCTGTGGCCTGT 60
QY 61 GGGAGAGGAAAAAGAGAGAGGACCGGCCCAAGCAAGCAAGCGAGGCCAGGCCGC 120


```
QY 421 GGCTGCTGATGGCGCTGCAGACAGAGCGCGCGGCTCGCGCGCGGCTCTCGCGTTG 480
Db 421 GGCTGCTGATGGCGCTGCAGACAGAGCGCGCGGCTCGCGCGCGGCTCTCGCGTTG 480
QY 481 CTCCAGCACGGCGCTCGTGGTGAACCTTGGGCTTGGGGGTGAGCGCGGAGCCACACTG 540
Db 481 CTCCAGCACGGCGCTCGTGGTGAACCTTGGGCTTGGGGGTGAGCGCGGAGCCACACTG 540
QY 541 CCGCTCCTACGACGCTGCAGAGGCTGGCGGGATCACGCTGGTGGATGTCTCGGCGG 600
Db 541 CCGCTCCTACGACGCTGCAGAGGCTGGCGGGATCACGCTGGTGGATGTCTCGGCGG 600
QY 601 GAGCAACTCCGGAGCGGCCAGACCGCGAGAGCGGAGAGCGCCATCGAGCGCGG 660
Db 601 GAGCAACTCCGGAGCGGCCAGACCGCGAGAGCGGAGAGCGCCATCGAGCGCGG 660
QY 661 GCACGGCGAGCTCAGCGATCTGGAGTCGGATCTGGCGGGCACAAAGACTGGCCCGT 720
Db 661 GCACGGCGAGCTCAGCGATCTGGAGTCGGATCTGGCGGGCACAAAGACTGGCCCGT 720
QY 721 ACCGGCGGCAACCGCGCTGCGGAGCTGATCGTCCGCGGAGAGAGAGAGATCCAGAGTT 780
Db 721 ACCGGCGGCAACCGCGCTGCGGAGCTGATCGTCCGCGGAGAGAGAGATCCAGAGTT 780
QY 781 CTTCCGCGCGCGAGCGGCCAGGCCAAGCGCTTGGCTTCCAAAGTACAACTTCGACTT 840
Db 781 CTTCCGCGCGCGAGCGGCCAGGCCAAGCGCTTGGCTTCCAAAGTACAACTTCGACTT 840
QY 841 CGTCCGCGCGCTGCCCTCGACCGCGCGCGGCTCGAGTGGCGCGCGTGGTGGT 900
Db 841 CGTCCGCGCGCTGCCCTCGACCGCGCGCGGCTCGAGTGGCGCGCGTGGTGGT 900
QY 901 CTGAAGCGAGCGTGGCTCGGTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 960
Db 901 CTGAAGCGAGCGTGGCTCGGTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 960
QY 961 CCCCCCCCCCAACAAATACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db 961 CCCCCCCCCCAACAAATACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1021 GCCTAAGCTAACCAACCACTTATCTCTGCTCAATGATGCTTCTTCTCTGGAGC 1080
Db 1021 GCCTAAGCTAACCAACCACTTATCTCTGCTCAATGATGCTTCTTCTCTGGAGC 1080
QY 1081 TAGCAGGAGCGTAGTATATTTAGTACTACTTTTACTTATTCAGAGGTTATCTTGACCCC 1140
Db 1081 TAGCAGGAGCGTAGTATATTTAGTACTACTTTTACTTATTCAGAGGTTATCTTGACCCC 1140
QY 1141 GATAGATCAATCCGCTTACTGTGTATTTCTCTCATGCTCTCTTAGATGGAGTTAATC 1200
Db 1141 GATAGATCAATCCGCTTACTGTGTATTTCTCTCATGCTCTCTTAGATGGAGTTAATC 1200
QY 1201 GTCTTAATTTACTGTACAGCAGCTTGTGTTGGCTTGCAAGAGAGATCTGGTTGTCT 1260
Db 1201 GTCTTAATTTACTGTACAGCAGCTTGTGTTGGCTTGCAAGAGAGATCTGGTTGTCT 1260
QY 1261 CAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 CAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 TACGCGTGCATGCGAGCTCATAGCTCTTCTATAGTGTCACTTAATTCATC 1372
Db 1321 TACGCGTGCATGCGAGCTCATAGCTCTTCTATAGTGTCACTTAATTCATC 1372
```

RESULT 3
US-10-425-114-18379/c
; Sequence 18379, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 18379
LENGTH: 985
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3079-036-F9_FLI
US-10-425-114-18379

Query Match 26.2%; Score 360; DB 13; Length 985;
Best Local Similarity 99.8%; Pred. No. 2.7e-164;
Matches 410; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 549 ACAGCGCTGCAGAGGCTGGCGGGGATCACGTCCTGTGGATGTCTCGCGCGGCAACT 608
Db 674 ACAGCGCTGCAGAGGCTGGCGGGGATCACGTCCTGTGGATGTCTCGCGCGGCAACT 615
QY 609 CCGGAGCGGCCAGACCGCGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 668
Db 614 CCGGAGCGGCCAGACCGCGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 555
QY 669 AGCTCAGCGATCTGGAGTCGGATCTGGCGGGGACAAAGACTGGCCCCGCTCGCTACCGCGG 728
Db 554 AGCTCAGCGATCTGGAGTCGGATCTGGCGGGGACAAAGACTGGCCCCGCTCGCTACCGCGG 495
QY 729 CAACGCCGGCTCGGAGCTGATCGTCCGCGGGGACAAAGACTCCAGAGATCTTCGCGCG 788
Db 494 CAACGCCGGCTCGGAGCTGATCGTCCGCGGGGACAAAGACTCCAGAGATCTTCGCGCG 435
QY 789 CCGCGAGCGGCCAGGCCAAGCGCTTTGCTTCCAAAGTACAACTTCGCTCGCGCG 848
Db 434 CCGCGAGCGGCCAGGCCAAGCGCTTTGCTTCCAAAGTACAACTTCGCTCGCGCG 375
QY 849 GCGTCCCTCGACCGCGCGCGGCTTCGAGTGGCGCGCGGTTGGTTCAGCATCTGAAGCG 908
Db 374 GCGTCCCTCGACCGCGCGCGGCTTCGAGTGGCGCGCGGTTGGTTCAGCATCTGAAGCG 315
QY 909 AGCGTGGCTCGGTCGAAGGTGAAGCTAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 959
Db 314 AGCGTGGCTCGGTCGAAGGTGAAGCTAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 264

RESULT 4
US-10-425-114-34947
; Sequence 34947, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 34947
LENGTH: 1111
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: UC-ZMFLM017355F05_FLI

```
US-10-425-114-34947
Query Match      26.2%; Score 360; DB 13; Length 1111;
Best Local Similarity 99.8%; Pred. No. 2.7e-164;
Matches 410; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 549 ACACGCTGCAGAGGCTGGCGGGATCACGTCCTGGTGGATGTCTCGGCGGCGAGCAACT 608
    |||
Db 438 ACACGCTGCAGAGGCTGGCGGGATCACGTCCTGGTGGATGTCTCGGCGGCGAGCAACT 497

QY 609 CCGGAGCGGCCACGACCGGAGAGGCGGAGAGACGACCGCCATCGAGCCGGGCGCACGGCG 668
    |||
Db 498 CCGGAGCGGCCACGACCGGAGAGGCGGAGAGACGACCGCCATCGAGCCGGGCGCACGGCG 557

QY 669 AGCTCAGCGATCTGGAGTCGGATCTGGCGGGGCAAGAACTGGCCCGTCGTCACCGCGG 728
    |||
Db 558 AGCTCAGCGATCTGGAGTCGGATCTGGCGGGGCAAGAACTGGCCCGTCGTCACCGCGG 617

QY 729 CAACGCCGGCTGCGGAGCTGATCGTGCCGCCAGCACAGATCCAGGAGTTCTTCGCCG 788
    |||
Db 618 CAACGCCGGCTGCGGAGCTGATCGTGCCGCCAGCACAGATCCAGGAGTTCTTCGCCG 677

QY 789 CCGCGAGCGGCCAGGCGGCTTGTCTTCAAGTACAACTTCGACTTCGTCGCCGCG 848
    |||
Db 678 CCGCGAGCGGCCAGGCGGCTTGTCTTCAAGTACAACTTCGACTTCGTCGCCGCG 737

QY 849 GCGTGCCCTCGACGCGCGCGCGGTTGAGTGGCGCGGTTGTCAGCATCTGAAGCG 908
    |||
Db 738 GCGTGCCCTCGACGCGCGCGCGGTTGAGTGGCGCGGTTGTCAGCATCTGAAGCG 797

QY 909 ACGTGCGTCCGGTGCAAGGTGAAGCTAGAAAGAGAGAAAGATGCCCCCCCC 959
    |||
Db 798 ACGTGCGTCCGGTGCAAGGTGAAGCTAGAAAGAGAGAAAGATGCCCCCCCC 848

RESULT 5
US-09-764-891-2570
; Sequence 2570, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2570
; LENGTH: 149
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-2570

Query Match      7.9%; Score 108; DB 10; Length 149;
Best Local Similarity 100.0%; Pred. No. 6.5e-42;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAGCTTACGT 1321
    |||
Db 12 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAGCTTACGT 71

QY 1322 ACGGTGTCATGCGACGTATAGTCTTCTATAGTGTACCTAAATTCA 1369
    |||
Db 72 ACGGTGTCATGCGACGTATAGTCTTCTATAGTGTACCTAAATTCA 119

RESULT 6
US-09-925-297-262
; Sequence 262, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 262
; LENGTH: 1557
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (9)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1347)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1527)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1533)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-297-262

Query Match      7.9%; Score 108; DB 9; Length 1557;
Best Local Similarity 100.0%; Pred. No. 5e-42;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAGCTTACGT 1321
    |||
Db 1211 AAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAGCTTACGT 1270

QY 1322 ACGGTGTCATGCGACGTATAGTCTTCTATAGTGTACCTAAATTCA 1369
    |||
Db 1271 ACGGTGTCATGCGACGTATAGTCTTCTATAGTGTACCTAAATTCA 1318

RESULT 7
US-10-023-896-42
; Sequence 42, Application US/10023896
; Publication No. US2003002776A1
; GENERAL INFORMATION:
; APPLICANT: Victor Roschke
; TITLE OF INVENTION: 29 Human Cancer Associated Proteins
; FILE REFERENCE: PA004P1
; CURRENT APPLICATION NUMBER: US/10/023,896
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: PCT/US00/23794
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/152,296
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/158,003
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
; LENGTH: 1557
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: n equals a,t,g, or c
```


NAME/KEY: misc feature
LOCATION: (9)..(9)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (1347)..(1347)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (1527)..(1527)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (1533)..(1533)
OTHER INFORMATION: n equals a,t,g, or c
US-10-023-896-42

Query Match 7.9%; Score 108; DB 15; Length 1557;
Best Local Similarity 100.0%; Pred. No. 5e-42;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGT 1321
Db 1211 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGT 1270
QY 1322 ACGGTGCATGCGACGTATAGCTCTTCTATAGTGTACCTAAATTCA 1369
Db 1271 ACGGTGCATGCGACGTATAGCTCTTCTATAGTGTACCTAAATTCA 1318

RESULT 8

US-09-925-297-261
Sequence 261, Application US/09925297
Patent No. US20020081659A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA105
CURRENT APPLICATION NUMBER: US/09/925,297
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05989
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 928
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 261
LENGTH: 2116
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (4)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (7)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (16)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (25)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (35)
OTHER INFORMATION: n equals a,t,g, or c
US-09-925-297-261

Query Match 7.9%; Score 108; DB 9; Length 2116;
Best Local Similarity 100.0%; Pred. No. 4.8e-42;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGT 1321
Db 1977 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGT 2036

QY 1322 ACGGTGCATGCGACGTATAGCTCTTCTATAGTGTACCTAAATTCA 1369
Db 2037 ACGGTGCATGCGACGTATAGCTCTTCTATAGTGTACCTAAATTCA 2084

RESULT 9

US-10-023-896-31
Sequence 31, Application US/10023896
Publication No. US20030027776A1
GENERAL INFORMATION:
APPLICANT: Victor Roschke
TITLE OF INVENTION: 29 Human Cancer Associated Proteins
FILE REFERENCE: PA004P1
CURRENT APPLICATION NUMBER: US/10/023,896
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: unassigned
PRIOR FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: PCT/US00/23794
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/152,296
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: 60/158,003
PRIOR FILING DATE: 1999-10-06
NUMBER OF SEQ ID NOS: 138
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 31
LENGTH: 2116
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (4)..(4)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (7)..(7)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (16)..(16)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (25)..(25)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (35)..(35)
OTHER INFORMATION: n equals a,t,g, or c
US-10-023-896-31

Query Match 7.9%; Score 108; DB 15; Length 2116;
Best Local Similarity 100.0%; Pred. No. 4.8e-42;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1262 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGT 1321
Db 1977 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCTCTAGAGGATCCAAAGCTTACGT 2036
QY 1322 ACGGTGCATGCGACGTATAGCTCTTCTATAGTGTACCTAAATTCA 1369
Db 2037 ACGGTGCATGCGACGTATAGCTCTTCTATAGTGTACCTAAATTCA 2084

RESULT 10

US-10-106-698-519
Sequence 519, Application US/10106698
Publication No. US20030109690A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
FILE REFERENCE: PA005P1
CURRENT APPLICATION NUMBER: US/10/106,698
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: PCT/US00/26524
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 60/157,137


```
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 519
; LENGTH: 701
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9)..(9)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (34)..(34)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (600)..(600)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (637)..(637)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (647)..(647)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (676)..(676)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (700)..(700)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-519
```

```
Query Match 7.7%; Score 106; DB 15; Length 701;
Best Local Similarity 100.0%; Pred. No. 5.1e-41;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1264 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAGCTTACGTAC 1323
Db 459 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAGCTTACGTAC 518

QY 1324 GCGTGCATGCGACGTATAGCTCTTCTATAGTGTCAACCTAAATTCA 1369
Db 519 GCGTGCATGCGACGTATAGCTCTTCTATAGTGTCAACCTAAATTCA 564
```

```
RESULT 11
US-10-264-049-133/c
; Sequence 133, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 133
; LENGTH: 845
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47)..(47)
; OTHER INFORMATION: n equals a,t,g, or c
```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (691)..(691)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (845)..(845)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-049-133
```

```
Query Match 7.5%; Score 103; DB 16; Length 845;
Best Local Similarity 100.0%; Pred. No. 1.4e-39;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1267 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAGCTTACGTACGCG 1326
Db 178 AAAAAAAAAAAAAAAAAAGGGCGGCTCTAGAGGATCCAAGCTTACGTACGCG 119

QY 1327 TGCATGCGACGTATAGCTCTTCTATAGTGTCAACCTAAATTCA 1369
Db 118 TGCATGCGACGTATAGCTCTTCTATAGTGTCAACCTAAATTCA 76
```

```
RESULT 12
US-10-264-237-1380/c
; Sequence 1380, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1380
; LENGTH: 2049
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (12)..(12)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (48)..(48)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (322)..(322)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1819)..(1819)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1987)..(1987)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2021)..(2021)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-1380
```

```
Query Match 7.2%; Score 99; DB 16; Length 2049;
Best Local Similarity 100.0%; Pred. No. 1.1e-37;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1271 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGCA 1330
Db 321 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGCA 262
QY 1331 TGGACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1369
Db 261 TGGACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 223

RESULT 13
US-10-050-704-66
; Sequence 66, Application US/10050704
; Publication No. US20030050442A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: PZ039P1
; CURRENT APPLICATION NUMBER: US/10/050,704
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: 09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 66
; LENGTH: 2657
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (179)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-050-704-66

Query Match 7.1%; Score 98; DB 15; Length 2657;
Best Local Similarity 100.0%; Pred. No. 3.4e-37;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1272 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGCA 1331
Db 2307 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGCA 2366
QY 1332 GCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1369
Db 2367 GCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 2404

RESULT 14
US-10-798-512-66
; Sequence 66, Application US/10798512
; Publication No. US20040152164A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 62 Human Secreted Proteins
; FILE REFERENCE: PZ039P1
; CURRENT APPLICATION NUMBER: US/10/798,512
; CURRENT FILING DATE: 2004-03-12
; PRIOR APPLICATION NUMBER: US/09/684,524
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/US00/08979
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,693
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/130,991
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 344
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 66

; LENGTH: 2657
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (179)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-798-512-66

Query Match 7.1%; Score 98; DB 17; Length 2657;
Best Local Similarity 100.0%; Pred. No. 3.4e-37;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1272 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGCA 1331
Db 2307 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGCA 2366
QY 1332 GCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 1369
Db 2367 GCGACGTCATAGCTCTTCTATAGTGTACCTAAATTCA 2404

RESULT 15
US-10-106-698-2395
; Sequence 2395, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 2395
; LENGTH: 544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (505)..(505)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (511)..(511)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (519)..(519)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (523)..(523)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (533)..(533)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-2395

Query Match 7.1%; Score 97; DB 15; Length 544;
Best Local Similarity 100.0%; Pred. No. 1.2e-36;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1273 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGCA 1332
Db 236 AAAAAAAAAAAAAAAAAAGGGCGCGCTCTAGAGGATCCAAAGCTTACGTACGCGTGCA 295

QY 1333 CGACGTCATAGCTCTTCTATAGTGTGCACCTAAATTCA 1369
|||
Db 296 CGACGTCATAGCTCTTCTATAGTGTGCACCTAAATTCA 332

Search completed: October 2, 2004, 10:11:41
Job time : 704 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 16:46:56 ; Search time 40 Seconds
(without alignments)
615.626 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 1304
Sequence: 1 MGKYMRCRGAAGAEVAAVE.....VRGVPLDAGGRFEWAPVVS I 256

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: Pirl1:*
2: Pirl2:*
3: Pirl3:*
4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|----------|--------------------|
| 1 | 178 | 13.7 | 196 | 2 T09968 | cyclin-dependent k |
| 2 | 149 | 11.4 | 327 | 2 T00797 | hypothetical prote |
| 3 | 129.5 | 9.9 | 1156 | 2 T34852 | probable secreted |
| 4 | 128 | 9.8 | 801 | 2 T29018 | hypothetical prote |
| 5 | 124 | 9.5 | 825 | 1 EDBEXD | immediate-early pr |
| 6 | 121.5 | 9.3 | 195 | 2 H96532 | hypothetical prote |
| 7 | 121.5 | 9.3 | 1151 | 2 T18535 | high molecular mas |
| 8 | 121.5 | 9.3 | 2639 | 2 T31328 | fibroin - Chinese |
| 9 | 121 | 9.3 | 191 | 2 T01132 | cyclin-dependent k |
| 10 | 115 | 8.8 | 1414 | 1 S23809 | collagen alpha 2(I |
| 11 | 113.5 | 8.7 | 467 | 2 E70976 | hypothetical prote |
| 12 | 113 | 8.7 | 1541 | 2 T02831 | AAA protein L4171. |
| 13 | 113 | 8.7 | 4391 | 2 A38096 | perlecan precursor |
| 14 | 112 | 8.6 | 1446 | 1 A45344 | immediate-early pr |
| 15 | 110.5 | 8.5 | 1077 | 2 A44067 | serine-rich protei |
| 16 | 109.5 | 8.4 | 294 | 2 S13141 | hypothetical prote |
| 17 | 107.5 | 8.2 | 490 | 2 T09084 | phosphatidylinosit |
| 18 | 107.5 | 8.2 | 698 | 2 T01209 | starch synthase (E |
| 19 | 107 | 8.2 | 278 | 2 S44796 | F09G8.6 protein - |
| 20 | 107 | 8.2 | 572 | 2 T08509 | trbL protein - Ent |
| 21 | 107 | 8.2 | 632 | 2 T00084 | hypothetical prote |
| 22 | 106.5 | 8.2 | 456 | 2 G97677 | hypothetical prote |
| 23 | 106 | 8.1 | 306 | 2 F84276 | forminoglutamate |
| 24 | 106 | 8.1 | 738 | 2 E87627 | hypothetical prote |
| 25 | 106 | 8.1 | 1147 | 2 T35781 | hypothetical prote |
| 26 | 105.5 | 8.1 | 1122 | 2 G64887 | probable tail fi |
| 27 | 105 | 8.1 | 237 | 2 A88640 | protein C34H4.4 [i |
| 28 | 105 | 8.1 | 583 | 1 S22544 | transcription fact |
| 29 | 104.5 | 8.0 | 413 | 2 AH2743 | conserved hypothet |

| | | | | | |
|----|-------|-----|------|----------|--------------------|
| 30 | 104.5 | 8.0 | 416 | 2 G97524 | hypothetical prote |
| 31 | 104.5 | 8.0 | 1487 | 1 EDBEE1 | immediate-early pr |
| 32 | 104.5 | 8.0 | 1487 | 1 EDBEF6 | 155K transcription |
| 33 | 104 | 8.0 | 220 | 2 A46597 | acidic calmodulin- |
| 34 | 104 | 8.0 | 397 | 2 A33880 | syndecan 2 - human |
| 35 | 104 | 8.0 | 888 | 2 I58378 | tyrosine kinase - |
| 36 | 103 | 7.9 | 381 | 2 S16506 | hypothetical prote |
| 37 | 102.5 | 7.9 | 340 | 2 T20807 | hypothetical prote |
| 38 | 102.5 | 7.9 | 730 | 2 A36226 | collagen alpha 1 c |
| 39 | 102.5 | 7.9 | 1174 | 2 A40853 | potassium channel |
| 40 | 102 | 7.8 | 502 | 2 C84400 | phenylalanyl-trNA |
| 41 | 102 | 7.8 | 571 | 2 T43456 | hypothetical prote |
| 42 | 101.5 | 7.8 | 411 | 2 G75475 | conserved hypothet |
| 43 | 101.5 | 7.8 | 1400 | 2 T31555 | hypothetical prote |
| 44 | 101 | 7.7 | 335 | 2 S08341 | myristylated alani |
| 45 | 101 | 7.7 | 401 | 1 QXBPL1 | hypothetical prote |

ALIGNMENTS

RESULT 1
T09968

cyclin-dependent kinase inhibitor protein - red goosefoot

C;Species: Chenopodium rubrum (red goosefoot)

C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000

C;Accession: T09968

R;Fountain, M.D.; Renz, A.; Beck, E.

submitted to the EMBL Data Library, November 1997

A;Description: A plant G1 cyclin-dependent kinase inhibitor gene from a photoautotrophic

A;Reference number: Z16910

A;Accession: T09968

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: mRNA

A;Residues: 1-196 <FOU>

A;Cross-references: EMBL:AJ002173

A;Experimental source: photoautotrophic cells derived from hypocotyl tissue

C;Genetics:

A;Gene: CDK11

C;Keywords: protein kinase inhibitor

| | | | | |
|-----------------------|--------|--|-------|------------------------------------|
| Query Match | 13.7% | Score 178; | DB 2; | Length 196; |
| Best Local Similarity | 27.5%; | Pred. No. 1.6e-05; | | |
| Matches | 61; | Conservative | 29; | Mismatches 102; Indels 30; Gaps 7; |
| QY | 33 | SAAAATGGVAKVAPRRKRAPAGEPAAAVSAGDGGGCIYHLRSMLFMAPPQPOPSVDSVP | 92 | |
| Db | 2 | AAAAT-----PTSSPAKKIKKVKSSYNIPQLRRRKNLSAPENFALET-- | 46 | |
| QY | 93 | TPVEAADGAAGQQAALAAAGLRSCSSTASSVNLGLGQGRGSHTCRS-YDAAEAGGDHVLV | 151 | |
| Db | 47 | TPLEVAADVVEEEE-----VANCSS--SEVITTARSDFFPPSCSSNYDQLSSSEFEVVK | 97 | |
| QY | 152 | DVSAASNSGSGPDRETRTPSSRAHGELESDLESLAGHKTGPSLPAATPAAELIVPPAH | 211 | |
| Db | 98 | DDGLGNRTADPEVESGEASSKQK---ESHRTAREATKLDQDDYPATKSTVQIKMPSDS | 154 | |
| QY | 212 | EIQEFFAAAEAAQAKFPASKYNFDFVRGVPLDAGGRFEWAPV | 253 | |
| Db | 155 | EIEEFFAAAEKDLQKRFSEKYNFDFIVKDVPLK--GRYDWWPI | 194 | |

RESULT 2
T00797

hypothetical protein At2g32710 [imported] - Arabidopsis thaliana

N;Alternate names: hypothetical protein F24L7.15

C;Species: Arabidopsis thaliana (mouse-ear cross)

C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 16-Feb-2001

C;Accession: T00797; E84736

R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul,

submitted to the EMBL Data Library, February 1998

A;Description: Arabidopsis thaliana chromosome II BAC F24L7 genomic sequence.

A;Reference number: Z14204

A;Accession: T00797
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-327 <ROU>
A;Cross-references: EMBL:AC003974; NID:g2914688; PID:g2914702
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: E84736
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-327 <STO>
A;Cross-references: GB:AE002093; NID:g2914702; PIDN:AAC04492.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g32710; F24L7.15
A;Map position: 2
A;Introns: 193/2
C;Superfamily: Arabidopsis thaliana hypothetical protein F24L7.15

Query Match 11.4%; Score 149; DB 2; Length 327;
Best Local Similarity 21.7%; Pred. No. 0.0027;
Matches 81; Conservative 34; Mismatches 89; Indels 170; Gaps 16;

QY 1 MGKYMRCR-----GAAG-----AEVAAVEV-----TQVGVTRRSRAAATGGV 40
Db 1 MGKIRKSKIDGAGAGAGGGGGGGGESSIALMDVSPSSSSGLVLRSLALQKQ 60
QY 41 AKVAPRRKRAPAGEPAAAVSAG-----GDGSCYIHLRSRMLFMAP----- 81
Db 61 QRCLLQKPSPPSLPPTSASPNPPSKQKMKKKQKQNMDCGS-YLQLRSRRLQKKEPIVVI 119
QY 82 -----PQPSVDSVPTPVEAADGAAGQQAALAGLSRCSSTASSV 123
Db 120 RSTKRKQQRNETCGRNPNRSLDSI-----RGDG-----SRSDSVSESV 161
QY 124 NLGLGGQSGSHTCRS-----YDAAEAGGDHVLVDVSAASNSGSGPDRETTSSRA 176
Db 162 VFGKDKDLISEINKDPTFGQNFDFLEE---EHT-----QSFNRTTRESFPCSLI 207
QY 177 H-----GELSDLESLA-GHKTGSLPAATPAEELIVPPAHEIQEFF 217
Db 208 RRPEIMTTPGSSTKLINCVSESNQREDSLSRSHRRRPTTP-----EMDEFF 253
QY 218 AAEEAAQAKRFASK-----YNEFDFVRG 239
Db 254 SGAEQEQKQKQFIEKYVFRFICSVLLVMSFQVLFSSGLVSLMVSNVSNFFRYNFDVNE 313
QY 240 VPLDAGGRFEWAPV 253
Db 314 QPLP--GRFEWTKV 325

RESULT 3
T34852
probable secreted protein - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C;Accession: T34852
R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, February 1999
A;Reference number: Z21559
A;Accession: T34852
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1156 <OLI>
A;Cross-references: EMBL:AL035478; PIDN:CAB36606.1; GSPDB:GN00070; SCOEDB:SC2G5.19
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SC2G5.19

Query Match 9.9%; Score 129.5; DB 2; Length 1156;
Best Local Similarity 24.1%; Pred. No. 0.22;
Matches 61; Conservative 35; Mismatches 122; Indels 35; Gaps 6;
QY 3 KYMRKCRGAAGAAVEVTQVGVTRRSRAAATGGVAKVAPRRKRAPAGEPAAAVSAG 62
Db 307 KAANAAGAAASARTAVQASSAHNAAR-RSAFAATAAAQAAATAGRAAALAYSAVA 365
QY 63 GDG-----GSCYIHLRSRMLFMAPPQPPQPSVDS-----VPTPVEAADGAAGQ 106
Db 366 RDASKTKAARLAAEGARNAAKARKAAQALAAQATATQAAAAAGISAAATARDSAAA 425
QY 107 AALAAAGLSRCSSTASSVNLGLGGQSGSHTCRSYDAAEAGGDHVLVDVSAASNSGSGPDRE 166
Db 426 AAVAQAQSGAAQSEAAVAAAAEAADAQAARATKAANR-----AQLANTAAASAAA 477
QY 167 RRETTSPSSRAHGE-----LSDLESLAG-----HKTPSLPAATPAEELIVPPAHEIQEF 216
Db 478 ARKAADSAAHAHAEKAAADAADAADAAGEADDYANKAKAWAADSVAALAAKAVDDARAV 537
QY 217 FAAAAEAAQAKRFA 229
Db 538 EAAAREAAEKLA 550

RESULT 4

T29018
hypothetical protein ZK84.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C;Accession: T29018
R;Kirsten, J.
submitted to the EMBL Data Library, April 1995
A;Description: The sequence of C. elegans cosmid ZK84.
A;Reference number: Z20553
A;Accession: T29018
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-801 <KIR>
A;Cross-references: EMBL:U23181; PIDN:AAC48204.1; GSPDB:GN00020; CESP:ZK84.1
A;Experimental source: strain Bristol N2; clone ZK84
C;Genetics:
A;Gene: CESP:ZK84.1
A;Map position: 2
A;Introns: 22/2; 45/3; 108/1
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;

Query Match 9.8%; Score 128; DB 2; Length 801;
Best Local Similarity 27.1%; Pred. No. 0.19;
Matches 67; Conservative 16; Mismatches 104; Indels 60; Gaps 9;
QY 11 AAGAEVAAVEVTQVGVTRRSRAAATGGVAKVAPRRKRAPAGE--PAAAVSAGDGGSC 68
Db 361 APATDIAATETTPAPSPAPVADAAAGYDSPSSIPEETPAPAAEDTPAPASAAEE--- 416
QY 69 YIHLRSRMLFMAPPQPPQPSVDSVPTP--VEAADGAAGQQAALAGLSRCSSTA---SSV 123
Db 417 -----TPAPAPAAETPAPETASAPDAAGGAAPADVAAPADVATTAPETSSA 464
QY 124 NLGLGGQSGSHTCRSYD-----AAEAGGDHVLVDVSAASNSGSGP----- 163
Db 465 QSAAG-----SYDVPSEPAEVTAPIVESATEAPSDSAAPIGPAASEPAPAPIEAP 515
QY 164 --DRERRETTSPSSRAHGEISDLSDLAGHKTGPSL-----PAATPAEELIVPPAHEIQEFF 217
Db 516 ATDAATLETAPAPAA--EPAPAAEAAAAGYDAPSSVPEETPAPAPAADETPAPAPAAEETP 573
QY 218 AAEEAAQ 224
Db 574 APAPAAE 580

A;Introns: 14/3

Query Match 9.3%; Score 121.5; DB 2; Length 2639;
Best Local Similarity 26.6%; Pred. No. 1.8;
Matches 75; Conservative 11; Mismatches 101; Indels 95; Gaps 10;

QY 11 AAGAEVAAVEVTQVGVVR-----TRSRSAATGGVAKVAPRR-----KRAPAG 53
Db 2010 AAAAAAAAAAGSGCGRGDGGYSGSSAAAAAAAAAAAAARRAGHDRAGSAAAAAAA 2069

QY 54 EPAAAVSAGDGG-----SCYIHLRSRLFMAPPQPSVDSVPTFVEAADGAAGQOG--- 106
Db 2070 AAAAAAGSGGSGYGGY-----GSDSAAAAAAAAAAAAAAGSGAGG 2112

QY 107 -----AALAAGLSRCSSSTASSVNLGLGGQSGSHTCRSYDAAEAGGDHV 149
Db 2113 AGGYGGYGGSDSAAAAAAAAAAGSGAGGCGYGGWGDGYSDSAAAAAAA--- 2169

QY 150 LVDVSAASNSGSGPDRERRETTPSSRAHGELSDLESDLAGHKTGPSLPAATPAAELIVPP 209
Db 2170 ---AAAAGSGAG-----GRGDG-----GYGSGSSAAAAAAAARR 2205

QY 210 A-HEIQEFAAAEAQAQKRFASKYNFDFVRGVPLDAGGRFEW 250
Db 2206 AGHERAAGSAAAAAAAAAASGAG-----RSGSGYGW 2238

RESULT 9
T01132
cyclin-dependent kinase inhibitor protein [imported] - Arabidopsis thaliana
N;Alternate names: hypothetical protein F26B6.8
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 16-Feb-2001
C;Accession: T01132; F84624
R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, submitted to the EMBL Data Library, June 1998
A;Description: Arabidopsis thaliana chromosome II BAC F26B6 genomic sequence.
A;Reference number: Z14198
A;Accession: T01132
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-191 <ROU>
A;Cross-references: EMBL:AC003040; NID:g3242700; PID:g3242706
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: F84624
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-191 <STO>
A;Cross-references: GB:AE002093; NID:g3242706; PIDN:AAC23758.1; GSPDB:GN00139
C;Genetics:
A;Gene: F26B6.8; At2g23430
A;Map position: 2
A;Introns: 66/3; 81/2; 170/2

Query Match 9.3%; Score 121; DB 2; Length 191;
Best Local Similarity 25.5%; Pred. No. 0.13;
Matches 59; Conservative 26; Mismatches 74; Indels 72; Gaps 12;

QY 46 RRKRAPAGEPAAVSAAGDGGSCYIHLRSRLFMAPPQPSVDSVPTFVEAADGAAGQ 105
Db 3 RKYRKAKGIVEAGVS-----STYMLRSRRIYVRSEKSSVSW-----GDN 45

QY 106 GAALAAGLSRCSSSTASSVNLGLGGQSGSHTCRS-----YDAAEAGGDHVLVDVSAASNSG 160
Db 46 GVS-----SSCS-----GSNEYKKKELIHLEEDKGD-----TETSTYR 80

QY 161 SGPDR-----ERRETTPSSRAHGELSDLESDL-----AGHKTGPSLPAA--TP 201

Db 81 RGTKRKLFFENLREEKEELSKSMENYS--SEFESAVKESLDCCCSGRKTMEETVTAESEE 138

QY 202 AAELI--VPPAHEIQEFAAAEAQAQKRFASKYNFDFVRGVPLDAGGRFEW 250
Db 139 KAKLMTMPTESEIEDFFVEAEKQLKEKFKKYNFDFEKEKPLE--GRYEW 187

RESULT 10
S23809
collagen alpha 2(I) chain homolog - sea urchin (Strongylocentrotus purpuratus)
C;Species: Strongylocentrotus purpuratus (purple urchin)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S23809
R;Exposito, J.Y.; d'Alessio, M.; Solursh, M.; Ramirez, F.
J. Biol. Chem. 267, 15559-15562, 1992
A;Title: Sea urchin collagen evolutionarily homologous to vertebrate pro-alpha-2(I) collagen
A;Reference number: S23809; MUID:92348411; PMID:1639795
A;Accession: S23809
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1414 <EXP>
A;Cross-references: EMBL:M92040; NID:g161435; PIDN:AAA30035.1; PID:g161436
C;Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homology
C;Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix
F;1207-1414/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match 8.8%; Score 115; DB 1; Length 1414;
Best Local Similarity 28.0%; Pred. No. 2.7;
Matches 60; Conservative 6; Mismatches 72; Indels 76; Gaps 8;

QY 10 GAAGAEVAAVEVTQVGVTRSRSAATGGVAKVAPRRKRAPAGEPAAAVSAGDGGSCY 69
Db 939 GPAGPPGAA-----GSRGPAGKSGDRGSPGAVGAGNPGPAGENGMPGSDGNDGA--- 988

QY 70 IHLRSRLFMAPPQPS-----VDSVPTFVEAADGAAGCQAALAAGLS-----RCSSTAS 121
Db 989 -----PGPQSGRGEKGDGTGASGANGSPGAPGPIGAPGAAGASGPRGETGTGA 1036

QY 122 SVNGLGGQSGSHTCRSYDAAEAGGDHVLVDVSAASNSG-SGPDRETRRTPSSRAHGEL 180
Db 1037 PGPLGPTGARGS-----TGPA GSPGSPGAPGERGTGP----- 1069

QY 181 SDLES DLAGHK-----TGPSLPAATPAA 203
Db 1070 -----AGKHGHPGVSGLPGLQGTSGPMGEPGA 1096

RESULT 11
E70976
hypothetical protein Rv3439c - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: E70976
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: E70976
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-467 <COL>
A;Cross-references: GB:Z95389; GB:AL123456; NID:g3242261; PIDN:CAB08688.1; PID:e316047;
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: Rv3439c

Query Match 8.7%; Score 113.5; DB 2; Length 467;
Best Local Similarity 27.4%; Pred. No. 1.1;
Matches 60; Conservative 16; Mismatches 68; Indels 75; Gaps 13;

QY 43 VAPRRKRAPAGEPA-----AAVSAGDGGSCYIHLRSRMLFMAPPQPPSVDSVP-TPVE 96
Db 230 LGPGRQSPASVPAQPSATAITPAA-----ALPPDPVPAVTSRPVTPSD 275
QY 97 ----AADGAAGQOGAALAAGLSRCSSTASSVNLGLGGQRG--SHTCRSYDA-----A 142
Db 276 FGSAPGDSATPAGVGSAGGFGDAGTG-----GLGGFAGLAGLANRIVDAVDSLLGSVA 330
QY 143 EAGGDHVLVD-----VSAASNSGSGPD-----RERRETPSSRAHGELS DL 183
Db 331 EQLGDLAADNPPGAVDPFAEDADNADDDGDDAHPEEADAEAEPEKEATEPDEA-DEVDDA 389
QY 184 E-----SDLAGHKGTGPSL-----PAATPAAELIVPP 209
Db 390 DESVPAERAQDVAEEATLPPVAEPPPPPAAPPVAE---PP 425
RESULT 12
T02831
AAA protein L4171.3 [imported] - Leishmania major (strain Friedlin)
C;Species: Leishmania major
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000
C;Accession: C81460; T02831
R;Myler, P.J.; Audleman, L.; devos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.;
Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
A;Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c
A;Reference number: A81455; MUID:99178987; PMID:10077609
A;Accession: C81460
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1541 <PYL>
A;Cross-references: GB:AE001274; NID:93264850; PIDN:AAC24654.1; PID:g2266909; GSPDB:GN00
A;Experimental source: strain MHOM/IL/81/Friedlin
C;Genetics:
A;Gene: L4171.3
A;Map position: 1

Query Match 8.7%; Score 113; DB 2; Length 1541;
Best Local Similarity 27.4%; Pred. No. 4;
Matches 66; Conservative 17; Mismatches 108; Indels 50; Gaps 9;
QY 11 AAGAEVAAVEVTQVGVTRRSRAAATGGVAKVAP---RRKRAPAGEPAAAVSAGDGGS 67
Db 1315 AAATTAAMVKTAASAPITTPAAQEAEEAKVDVVLPSVKARKKAAKEKAV----- 1365
QY 68 CYIHLRSRMLFMAPP---QPQPSVDSVPTPVEAADGAAGQOGAAL--AAGLSRC--SSTA 120
Db 1366 -----GVPPATGATPSTDAAPAVAAAKTTGGGKGKAVKDTGGGSRAPASVP 1414
QY 121 SSVNLGLGGQSGHTRSDAAEAGGDHVLVDVSAAS-----NSGSGPDRRR---ETTPS 173
Db 1415 SSPTSGKGG-----GTAHGKGTNDVKVSSSLDALMSASPPDAAEAALSQAPT 1464
QY 174 SRAHGELSDLES DLAGHKTPSL---PAATPAAELIVPPAHEIQEFAAAEAQAARFAS 230
Db 1465 SEKSTKSVFPQVDVAPALAAPAVAEAPTAKPADSKVAEPPEEAAVAPAAPAMATAVQDPS 1524
QY 231 K 231
Db 1525 K 1525

RESULT 13
A38096
perlecan precursor - human
N;Alternate names: basement membrane heparan sulfate proteoglycan; heparan sulfate prote
C;Species: Homo sapiens (man)
C;Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 05-Nov-1999
C;Accession: A38096; S19256; S77946; A41059; A40306; B33625; A33625; A41736
R;Murdoch, A.D.; Dodge, G.R.; Cohen, I.; Tuan, R.S.; Iozzo, R.V.
J. Biol. Chem. 267, 8544-8557, 1992
A;Title: Primary structure of the human heparan sulfate proteoglycan from basement memb

tor, laminin, neural cell adhesion molecules, and epidermal growth factor.
A;Reference number: A38096; MUID:92235084; PMID:1569102
A;Accession: A38096
A;Molecule type: mRNA
A;Residues: 1-4391 <MUR>
A;Cross-references: GB:M85289; NID:g184426; PIDN:AAA52700.1; PID:g184427
R;Kallunki, P.; Tryggvason, K.
J. Cell Biol. 116, 559-571, 1992
A;Title: Human basement membrane heparan sulfate proteoglycan core protein: a 467-kD prot
ell adhesion molecules, and epidermal growth factor.
A;Reference number: A41736; MUID:92112994; PMID:1730768
A;Accession: S19256
A;Molecule type: mRNA
A;Residues: 1-57,'D',59-434,'A',436,'FL',438-449,'Q',451-502,'A',503-792,'K',794-908,'R',
71-2979,'H',2981-2994,'G',2996-3167,'T',3169-3240,'R',3242-3426,'R',3428-3631,'Q',3633-4
A;Cross-references: EMBL:X62515; NID:g29469; PIDN:CAA44373.1; PID:g29470
R;Kallunki, P.; Eddy, R.L.; Byers, M.G.; Kestilae, M.; Shows, T.B.; Tryggvason, K.
Genomics 11, 389-396, 1991
A;Title: Cloning of human heparan sulfate proteoglycan core protein, assignment of the ge
A;Reference number: A41059; MUID:92120660; PMID:1685141
A;Accession: A41059
A;Molecule type: mRNA
A;Residues: 'RT',892-908,'R',910-1101,'L',1103-1132,'L',1134-1221,'L',1223-1397 <KA2>
A;Cross-references: GB:S76436; NID:g243370; PIDN:AAB21121.1; PID:g243371
R;Dodge, G.R.; Kovalszky, I.; Chu, M.L.; Hassell, J.R.; McBride, O.W.; Yi, H.F.; Iozzo, R
Genomics 10, 673-680, 1991
A;Title: Heparan sulfate proteoglycan of human colon: partial molecular cloning, cellular
A;Reference number: A40306; MUID:91365376; PMID:1679749
A;Accession: A40306
A;Molecule type: mRNA
A;Residues: 1018-1405,'G',1407-1409,'G',1411-1465 <DOD>
A;Cross-references: GB:M64283; NID:g184424; PIDN:AAA52699.1; PID:g184425
R;Heremans, A.; van der Schueren, B.; De Cock, B.; Paulsson, M.; Cassiman, J.J.; van den
J. Cell Biol. 109, 3199-3211, 1989
A;Title: Matrix-associated heparan sulfate proteoglycan: core protein-specific monoclonal
anes.
A;Reference number: A33625; MUID:90078352; PMID:2687294
A;Accession: B33625
A;Molecule type: protein
A;Residues: 1379-1384,'X',1386-1388,'X',1390-1398 <HE2>
A;Accession: A33625
A;Molecule type: protein
A;Residues: 2166-2171,'X',2173-2175,'X',2177-2185 <HE3>
A;Note: peptide potentially matches four different regions of sequence shown
C;Genetics:
A;Gene: GDB:HSPG2
A;Cross-references: GDB:126372; OMIM:142461
A;Map position: lp36.1-lp36.1
C;Superfamily: LDL receptor ligand-binding repeat homology; EGF homology; laminin G repea
C;Keywords: chondroitin sulfate proteoglycan; glycoprotein; heparan sulfate; transmembrar
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-4391/Product: perlecan #status predicted <MAT>
F;22-193/Domain: I <DOM1>
F;194-530/Domain: II <DOM2>
F;199-234/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F;285-319/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F;325-359/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F;368-403/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F;531-1676/Domain: III <DOM3>
F;1159-1206/Domain: laminin-type EGF-like homology <LEG>
F;1563-1610/Domain: laminin-type EGF-like homology <EG7>
F;1613-1668/Domain: laminin-type EGF-like homology <LEG8>
F;1677-3686/Domain: IV <DOM4>
F;2007-2034/Domain: transmembrane #status predicted <TRM>
F;3687-4391/Domain: V <DOM5>

F;3845-3880/Domain: EGF homology <EGF1>
F;3888-3921/Domain: EGF homology <EGF>
F;3953-4106/Domain: laminin G repeat homology <LG2>
F;4147-4175/Domain: EGF homology <EGF2>
F;4149-4151/Region: motor neuron attachment (L-R-E) motif
F;4299-4301/Region: motor neuron attachment (L-R-E) motif
F;65,71,76/Binding site: heparan sulfate (Ser) (covalent) #status predicted
F;89,554,1755,2121,3072,3105,3279,3780,3836,4068/Binding site: carbohydrate (Asn) (coval) #status predicted
F;2995,3933,4179/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

Query Match 8.7%; Score 113; DB 2; Length 4391;
Best Local Similarity 26.5%; Pred. No. 12;
Matches 75; Conservative 27; Mismatches 83; Indels 98; Gaps 16;

QY 2 GKYMRCRGAAGAEVAVET-----QVVG--VTRSRSAATGG-----VA 41
Db 2499 GEYVCRVVGSGTQEAASVLVTIQRLSGSHSQGVAVPVRIESSASLANGLTDLNCLVA 2558

QY 42 KVAPR-----RKRAP-----AGEPAAAVSAG----- 62
Db 2559 SQAPHTITWYKRGSLPSRHQIVGSRIRIPQVTPADSGEYVCHVSNAGSRETSLIVTIQ 2618

QY 63 GDGGSCYIHLRSMLEFMAPQPPQPSVDSVPTPVEA-----ADGAAGQQAALAGLSRCS 118
Db 2619 GSGSS---HVPS---VSPPIRES---SSPTVVEGOTLDLNCWARQPQAIIITWYKRGGS 2669

QY 119 TAS-----SVNLGLG---GQGSHTCR---SYDAEAGGDHVLVDVSAASNSGSGPDRE 166
Db 2670 LPSRHQTHGSHLRHLHQMVSADSGEYVCRANNIDALEAS---IVISVSPSAGSPSPAGSS 2726

QY 167 ---RRETPSSRAHGELSDLESLAG-----HKTGPSLPA 198
Db 2727 MPRIESSSHVAEGETLDLNCVVPQGAHAQVTHWKRGSLS 2769

RESULT 14
A45344
immediate-early protein - suid herpesvirus 1 (strain Kaplan)
C;Species: suid herpesvirus 1
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jul-1999
C;Accession: A45344
R;Vlcek, C.; Kozmik, Z.; Paces, V.; Schirm, S.; Schwyzler, M.
Virology 179, 365-377, 1990
A;Title: Pseudorabies virus immediate-early gene overlaps with an oppositely oriented op
A;Reference number: A45344; MUID:91021039; PMID:2171211
A;Accession: A45344
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-1446 <VLC>
A;Cross-references: GB:M34651; NID:g334070; PIDN:AAA47470.1; PID:g334071
C;Superfamily: herpesvirus immediate-early protein IE175
C;Keywords: DNA binding; early protein; transcription regulation

Query Match 8.6%; Score 112; DB 1; Length 1446;
Best Local Similarity 26.8%; Pred. No. 4.4;
Matches 64; Conservative 19; Mismatches 78; Indels 78; Gaps 11;

QY 9 RGAAGAEVAAVEVT-----QVVGVRTRSR-----AAATGGVAKVAPRRKRAPAGEPAA 58
Db 745 RRQADSVALVARTVAPLVRYSDGARAREAAWTYAAALFAPANVAARLAEAAARPQPA 804

QY 59 VSAGDGGSCYIHLRSMLEFMAPQPPQPSVDSVPTPVEADGA-AGQGAALAGLSRCS 117
Db 805 EPAPG-----LPPLWPEQPLVVPAPAPAAAGAPSLPGSGPSPASTKS 849

QY 118 S-----TASSVNLGLGGQGRGSHTCRSYDAEAGGDHVLVDVSAASNSGSGPD-----RERR 168
Db 850 SSSTKSSSSTKSLSGSSG-----YASSPAAGPDPAPERRKKK 887

QY 169 ETPSSRAHGELSD-----LESDLAHLK-----TGP-----SL-----PAATPAELI 206
Db 888 RRAPGARPPGDGEDEGLSGAALRGDGHGRDDEEDRGRPRKRRLGLGPADPAPALL 946

RESULT 15

A44067

serine-rich protein hairless - fruit fly (Drosophila melanogaster)
N;Alternate names: 109K basic protein H
C;Species: Drosophila melanogaster
C;Date: 10-Jun-1993 #sequence_revision 26-Feb-1999 #text_change 26-Feb-1999
C;Accession: A44067; A58929; S33412; S24639
R;Bang, A.G.; Posakony, J.W.
Genes Dev. 6, 1752-1769, 1992

A;Title: The Drosophila gene Hairless encodes a novel basic protein that controls alterna
A;Reference number: A44067; MUID:92387549; PMID:1516831
A;Accession: A44067
A;Molecule type: DNA
A;Residues: 19-1077 <BAN>
A;Cross-references: GB:M95192; NID:g157621; PID:g157622
A;Note: sequence extracted from NCBI backbone (NCBIN:112622, NCBIP:112623)
R;Preiss, A.

submitted to the EMBL Data Library, May 1994
A;Description: Hairless, a Drosophila gene involved in neural development, encodes a nove
A;Reference number: A58929
A;Accession: A58929

A;Molecule type: mRNA
A;Residues: 1-1077 <PRE>
A;Cross-references: EMBL:X67239; GB:S49642; NID:g578331; PID:g578332
R;Maier, D.; Stumm, G.; Kuhn, K.; Preiss, A.
Mech. Dev. 38, 143-156, 1992

A;Title: Hairless, a Drosophila gene involved in neural development, encodes a novel, sei
A;Reference number: S33412; MUID:93041287; PMID:1419850
A;Accession: S33412

A;Molecule type: mRNA
A;Residues: 1-150, 'A', 152-701, 'LL', 704-890, 'R', 892-963, 'RLLP', 968-973, 975-1077 <MAI>
A;Cross-references: EMBL:X67239
C;Genetics:

A;Gene: FlyBase:H; hairless
A;Cross-references: FlyBase:FBgn0001169

Query Match 8.5%; Score 110.5; DB 2; Length 1077;
Best Local Similarity 20.9%; Pred. No. 4.1;
Matches 58; Conservative 33; Mismatches 87; Indels 99; Gaps 11;

QY 33 SAAATGGVAK-----VAPRRK-----RAPAGEPAAAVSAGDGGSCYIHLRSR 75
Db 557 SATAAGRLVEYHTQHVSPRKRLREFEKFVSLDNGCVNNGSGGASGGAGGK-----RSR 611

QY 76 M-----LFMAPPPQPPSVDSVPTPVEAADGAAGQQAALAGLSRCSSTAS 121
Db 612 AKGTSTSSPAGKASPMNLAPPQPKPS---PSP---GSSSSTSPATLSTQPTRLNSSYS 664

QY 122 SVNGLGGQGRGSHTCRSYDAEAGGDHVLVDVS----- 154
Db 665 IHSI-LGGSSSGSGSSSSSSSGKCGDHPAAIISNVHHPQHSMYQPSSSSYPRALLTSPKS 723

QY 155 ---AASNSGSGPDRERRET-----TPSSRAHGELSD-----LESDLA 188
Db 724 PDVSGSNGGGKSPSHTGTKRSPPPYSAGSPVDYGHGFYRDPYAGAGRPSTSGSASQDLS 783

QY 189 GHKTGPSLPAATPAAEELIVPPAHEIQEFAAAEAQA 225
Db 784 PPRSSPASPATTP--RTVPKKTASIRREFASPSASS 818

Search completed: October 1, 2004, 16:54:33
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 16:45:15 ; Search time 24 Seconds
(without alignments)
555.415 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 1304
Sequence: 1 MGKYMRCRGAAGAEVAAVE.....VRGVPLDAGRFEPVWSI 256

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 124 | 9.5 | 825 | 1 | ICP0 HSV2H |
| 2 | 113 | 8.7 | 4391 | 1 | PGBM_HUMAN |
| 3 | 112 | 8.6 | 1446 | 1 | IE18_PRVKA |
| 4 | 111.5 | 8.6 | 1115 | 1 | TBC2_CHLRE |
| 5 | 110.5 | 8.5 | 1077 | 1 | HLES_DROME |
| 6 | 107 | 8.2 | 278 | 1 | YLS6_CAEEL |
| 7 | 106 | 8.1 | 306 | 1 | HUTC_HALN1 |
| 8 | 106 | 8.1 | 677 | 1 | OGFR_HUMAN |
| 9 | 106 | 8.1 | 1461 | 1 | IE18_PRVIF |
| 10 | 105.5 | 8.1 | 1120 | 1 | STFR_ECOLI |
| 11 | 105 | 8.1 | 553 | 1 | FXC1_MOUSE |
| 12 | 104.5 | 8.0 | 1487 | 1 | ICP4_HSVEB |
| 13 | 104.5 | 8.0 | 1487 | 1 | ICP4_HSVBK |
| 14 | 104 | 8.0 | 219 | 1 | BASP_RAT |
| 15 | 104 | 8.0 | 888 | 1 | KLTK_MOUSE |
| 16 | 103.5 | 7.9 | 388 | 1 | HXAD_HUMAN |
| 17 | 103.5 | 7.9 | 391 | 1 | SOX1_MOUSE |
| 18 | 103.5 | 7.9 | 518 | 1 | TPM4_DROME |
| 19 | 102.5 | 7.9 | 386 | 1 | HXAD_MOUSE |
| 20 | 102.5 | 7.9 | 553 | 1 | FXC1_HUMAN |
| 21 | 102.5 | 7.9 | 1174 | 1 | CIKE_DROME |
| 22 | 102 | 7.8 | 502 | 1 | SYFA_HALN1 |
| 23 | 101.5 | 7.8 | 447 | 1 | ODP2_RHIME |
| 24 | 101.5 | 7.8 | 704 | 1 | CLI6_HUMAN |
| 25 | 101 | 7.7 | 774 | 1 | STF_LAMBD |
| 26 | 100.5 | 7.7 | 864 | 1 | KLTK_HUMAN |
| 27 | 100.5 | 7.7 | 1093 | 1 | AF17_HUMAN |
| 28 | 100.5 | 7.7 | 1163 | 1 | Y222_HUMAN |
| 29 | 100 | 7.7 | 907 | 1 | AI180_HUMAN |
| 30 | 100 | 7.7 | 1321 | 1 | IRS2_MOUSE |
| 31 | 99.5 | 7.6 | 328 | 1 | HXD1_HUMAN |
| 32 | 99.5 | 7.6 | 775 | 1 | ICP0_HSV11 |
| 33 | 99.5 | 7.6 | 2333 | 1 | PGCA_CANFA |
| 1 | 124 | 9.5 | 825 | 1 | P28284 herpes simp |
| 2 | 113 | 8.7 | 4391 | 1 | P98160 homo sapien |
| 3 | 112 | 8.6 | 1446 | 1 | P33479 pseudorabie |
| 4 | 111.5 | 8.6 | 1115 | 1 | Q8vxp3 chlamydomon |
| 5 | 110.5 | 8.5 | 1077 | 1 | Q02308 drosophila |
| 6 | 107 | 8.2 | 278 | 1 | P34391 caenorhabdi |
| 7 | 106 | 8.1 | 306 | 1 | Q9hqd7 halobacteri |
| 8 | 106 | 8.1 | 677 | 1 | Q9nzt2 homo sapien |
| 9 | 106 | 8.1 | 1461 | 1 | P11675 pseudorabie |
| 10 | 105.5 | 8.1 | 1120 | 1 | P76072 escherichia |
| 11 | 105 | 8.1 | 553 | 1 | Q61572 mus musculu |
| 12 | 104.5 | 8.0 | 1487 | 1 | P28925 equine herp |
| 13 | 104.5 | 8.0 | 1487 | 1 | P17473 equine herp |
| 14 | 104 | 8.0 | 219 | 1 | Q05175 rattus norv |
| 15 | 104 | 8.0 | 888 | 1 | P08923 mus musculu |
| 16 | 103.5 | 7.9 | 388 | 1 | P31271 homo sapien |
| 17 | 103.5 | 7.9 | 391 | 1 | P53783 mus musculu |
| 18 | 103.5 | 7.9 | 518 | 1 | P49455 drosophila |
| 19 | 102.5 | 7.9 | 386 | 1 | Q62424 mus musculu |
| 20 | 102.5 | 7.9 | 553 | 1 | Q12948 homo sapien |
| 21 | 102.5 | 7.9 | 1174 | 1 | Q02280 drosophila |
| 22 | 102 | 7.8 | 502 | 1 | Q9hmk4 halobacteri |
| 23 | 101.5 | 7.8 | 447 | 1 | Q9r9n3 rhizobium m |
| 24 | 101.5 | 7.8 | 704 | 1 | Q96ny7 homo sapien |
| 25 | 101 | 7.7 | 774 | 1 | P03764 bacterioph |
| 26 | 100.5 | 7.7 | 864 | 1 | P29376 homo sapien |
| 27 | 100.5 | 7.7 | 1093 | 1 | P55198 homo sapien |
| 28 | 100.5 | 7.7 | 1163 | 1 | Q92618 homo sapien |
| 29 | 100 | 7.7 | 907 | 1 | O60641 homo sapien |
| 30 | 100 | 7.7 | 1321 | 1 | P81122 mus musculu |
| 31 | 99.5 | 7.6 | 328 | 1 | Q9gz00 homo sapien |
| 32 | 99.5 | 7.6 | 775 | 1 | P08393 herpes simp |
| 33 | 99.5 | 7.6 | 2333 | 1 | Q28343 canis famil |

| | | | | | | |
|----|------|-----|------|---|-------------|--------------------|
| 34 | 99 | 7.6 | 676 | 1 | ICP0 HSVBJ | P29128 bovine herp |
| 35 | 99 | 7.6 | 1186 | 1 | HCN4_MOUSE | O70507 mus musculu |
| 36 | 99 | 7.6 | 3707 | 1 | PGBM_MOUSE | Q05793 mus musculu |
| 37 | 98.5 | 7.6 | 901 | 1 | AI180_MOUSE | Q61548 mus musculu |
| 38 | 98.5 | 7.6 | 910 | 1 | PERT_BORPE | P14283 bordetella |
| 39 | 98 | 7.5 | 440 | 1 | DCO_DROME | O76324 drosophila |
| 40 | 98 | 7.5 | 1324 | 1 | IRS2_HUMAN | Q9y4h2 homo sapien |
| 41 | 97.5 | 7.5 | 676 | 1 | ICP0 HSVBK | P29836 bovine herp |
| 42 | 97.5 | 7.5 | 1733 | 1 | VNUA_PRVKA | P33485 pseudorabie |
| 43 | 97 | 7.4 | 2038 | 1 | FSH_DROME | P13709 drosophila |
| 44 | 96.5 | 7.4 | 336 | 1 | TTC9_HUMAN | Q92623 homo sapien |
| 45 | 96.5 | 7.4 | 421 | 1 | CCG8_RAT | Q8vhw5 rattus norv |

ALIGNMENTS

RESULT 1
ICP0 HSV2H
ID ICP0 HSV2H STANDARD; PRT; 825 AA.
AC P28284;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Trans-acting transcriptional protein ICP0 (VMW118 protein).
GN RL2.
OS Herpes simplex virus (type 2 / strain HG52).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10315;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92113549; PubMed=1662697;
RA McGeoch D.J., Cunningham C., McIntyre G., Dolan A.;
RT "Comparative sequence analysis of the long repeat regions and
RT adjoining parts of the long unique regions in the genomes of herpes
RT simplex viruses types 1 and 2.";
RL J. Gen. Virol. 72:3057-3075(1991).
RN [2]
RP SEQUENCE FROM N.A.

RA Dolan A.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -!- SIMILARITY: TO OTHER HERPESVIRUSES ICP0 PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D10471; BAA23427.1; -.
DR EMBL; Z86099; CAB06760.1; -.
DR PIR; JQ1501; EDBEXD.
DR HSSP; P28990; 1CHC.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SMO0184; RING; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Transcription regulation; Trans-acting factor; Activator; Zinc-finger;
DNA-binding.
DOMAIN 120 123 POLY-GLY.
ZN_FING 126 167 RING-TYPE.
DOMAIN 266 271 POLY-GLY.
DOMAIN 292 295 POLY-SER.
DOMAIN 342 345 POLY-ALA.
DOMAIN 386 389 POLY-SER.
DOMAIN 395 400 POLY-GLY.
DOMAIN 425 428 POLY-ALA.
DOMAIN 590 627 POLY-SER.

SQ SEQUENCE 825 AA; 81986 MW; 5CEB15858553A274 CRC64;
Query Match 9.5%; Score 124; DB 1; Length 825;
Best Local Similarity 28.3%; Pred. No. 0.4;
Matches 60; Conservative 18; Mismatches 82; Indels 52; Gaps 9;
QY 6 RKCGRGAAGAEVAEVTQVGVTRSRSAATGGVAKVAPRRKRAPAGE-----PAAA 58
Db 513 RKRRGSDSGPAASSASS--SAAPRSLAPQGVGAKRAAPR--RAPDSDSGRGHGLAP 568
QY 59 VSAGDGGSCYIHLRSMLFMAPPQPPQPSVDSVPTPVEAADGAAGQGAALAAGLSRCSS 118
Db 569 ASAGA-----APPSASPS-----SQAAVAASSSSSASSSSSAS 605
QY 119 TASSVNLGLGGQGSHTCRSYDAEAGGDHVLVDVSAASNSGSGPDRERETTPSPRAHG 178
Db 606 SSSASSSSASSSSASSSSSSAGGAGG-----SVASASGAG---ERRETSIGPRAAA 655
QY 179 ELSDLESDLGHKT-----GPSLPAATPAEL 205
Db 656 PRGPRK---CARKTRHAEGGPEPGARDPAPGL 684

RESULT 2

PGBM HUMAN STANDARD; PRT; 4391 AA.
AC P98160; Q16287; Q9H3V5;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Basement membrane-specific heparan sulfate proteoglycan core
DE protein precursor (HSPG) (Perlecan) (PLC).
CN HSPG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1] -
RP SEQUENCE FROM N.A.
RX MEDLINE=92112994; PubMed=1730768;
RA Kallunki P., Tryggvason K.;
RT "Human basement membrane heparan sulfate proteoglycan core protein: a
RT 467-kD protein containing multiple domains resembling elements of the
RT low density lipoprotein receptor, laminin, neural cell adhesion
RT molecules, and epidermal growth factor.";
RL J. Cell Biol. 116:559-571 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon, and Skin;
RX MEDLINE=92235084; PubMed=1569102;
RA Murdoch A.D., Dodge G.R., Cohen I., Tuan R.S., Iozzo R.V.;
RT "Primary structure of the human heparan sulfate proteoglycan from
RT basement membrane (HSPG2/perlecan). A chimeric molecule with multiple
RT domains homologous to the low density lipoprotein receptor, laminin,
RT neural cell adhesion molecules, and epidermal growth factor.";
RL J. Biol. Chem. 267:8544-8557 (1992).
RN [3]
RP SEQUENCE OF 22-4391 FROM N.A., AND VARIANT SJS1 TYR-1532.
RX MEDLINE=20553141; PubMed=1101850;
RA Nicole S., Davoine C.-S., Topaloglu H., Cattolico L., Barral D.,
RA Beighton P., Ben-Hamida C., Hammouda H., Cruaud C., White P.S.,
RA Samson D., Urtizberea J.A., Lehmann-Horn F., Weissenbach J.,
RA Hentati F., Fontaine B.;
RT "Perlecan, the major proteoglycan of basement membranes, is altered in
RT patients with Schwartz-Jampel syndrome (chondrodystrophic myotonia).";
RL Nat. Genet. 26:480-483 (2000).
RN [4]
RP SEQUENCE OF 1016-1470 FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=91365376; PubMed=1679749;
RA Dodge G.R., Kovalszky I., Chu M.L., Hassell J.R., McBride O.W.,
RA Yi H.F., Iozzo R.V.;
RT "Heparan sulfate proteoglycan of human colon: partial molecular

RT cloning, cellular expression, and mapping of the gene (HSPG2) to the
RT short arm of human chromosome 1.";
RL Genomics 10:673-680 (1991).
RN [5]
RP SEQUENCE OF 890-1396 FROM N.A.
RC TISSUE=Fibrosarcoma;
RX MEDLINE=92120660; PubMed=1685141;
RA Kallunki P., Eddy R.L., Byers M.G., Kestila M., Shows T.B.,
RA Tryggvason K.;
RT "Cloning of human heparan sulfate proteoglycan core protein,
RT assignment of the gene (HSPG2) to lp36.1-->p35 and identification of
RT a BamHI restriction fragment length polymorphism.";
RL Genomics 11:389-396 (1991).
RN [6]
RP SEQUENCE OF 1-21 FROM N.A.
RX MEDLINE=94052171; PubMed=8234307;
RA Cohen I.R., Graessel S., Murdoch A.D., Iozzo R.V.;
RT "Structural characterization of the complete human perlecan gene and
RT its promoter.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:10404-10408 (1993).
RN [7]
RP CARBOHYDRATE-LINKAGE SITE ASN-2121.
RX MEDLINE=22660472; PubMed=12754519;
RA Zhang H., Li X.-J., Martin D.B., Aebersold R.;
RT "Identification and quantification of N-linked glycoproteins using
RT hydrazide chemistry, stable isotope labeling and mass spectrometry.";
RL Nat. Biotechnol. 21:660-666 (2003).
CC -!- FUNCTION: This protein is an integral component of basement
CC membranes. It is responsible for the fixed negative electrostatic
CC charge and is involved in the charge-selective ultrafiltration
CC properties. It serves as an attachment substrate for cells.
CC -!- SUBUNIT: Purified perlecan has a strong tendency to aggregate in
CC dimers or stellate structures. It interacts with other basement
CC membrane components such as laminin, prolargin and collagen type
CC IV.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: Found in the basement membranes.
CC -!- PTM: CONTAINS THREE HEPARAN SULFATE CHAINS AS WELL AS N-LINKED
CC AND O-LINKED OLIGOSACCHARIDES.
CC -!- DISEASE: Defects in HSPG2 are the cause of Schwartz-Jampel
CC syndrome (SJS1) [MIM:255800]; a rare autosomal recessive disorder
CC characterized by permanent myotonia (prolonged failure of muscle
CC relaxation) and skeletal dysplasia, resulting in reduced stature,
CC kyphoscoliosis, bowing of the diaphyses and irregular epiphyses.
CC -!- SIMILARITY: Contains 4 LDL-receptor class A domains.
CC -!- SIMILARITY: Contains 11 laminin EGF-like domains.
CC -!- SIMILARITY: Contains 3 laminin IV domains.
CC -!- SIMILARITY: Contains 22 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 3 laminin G-like domains.
CC -!- SIMILARITY: Contains 4 EGF-like domains.
CC -!- SIMILARITY: Contains 1 SEA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X62515; CAA44373.1; -.
CC EMBL; M85289; AAA52700.1; -.
CC EMBL; AL445795; CAC18534.1; -.
CC EMBL; M64283; AAA52699.1; -.
CC EMBL; S76436; AAB21121.2; -.
CC EMBL; L22078; -; NOT_ANNOTATED_CDS.
CC PIR; A38096; A38096.
CC HSSP; P00740; 1EDM.
CC Siena-2DPAGE; P98160; -.
CC Genew; HGNC:5273; HSPG2.
CC MIM; 142461; -.
CC MIM; 255800; -.
CC InterPro; IPR008985; ConA_like_lec_g1.

DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003596; Ig_v.
DR InterPro; IPR000034; Laminin B.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000082; SEA_domain.
DR Pfam; PF00008; EGF; 4.
DR Pfam; PF00047; ig; 22.
DR Pfam; PF00052; laminin_B; 3.
DR Pfam; PF00053; laminin_EGF; 7.
DR Pfam; PF00054; laminin_G; 3.
DR Pfam; PF00057; ldl_recept_a; 4.
DR Pfam; PF01390; SEA; 1.
DR PRINTS; PR00261; LDLRECEPTOR.
DR ProDom; PD003031; Laminin_B; 3.
DR SMART; SM00181; EGF; 15.
DR SMART; SM00180; EGF_Lam; 12.
DR SMART; SM00409; IG; 22.
DR SMART; SM00408; IGC2; 21.
DR SMART; SM00406; IGV; 7.
DR SMART; SM00281; LamB; 3.
DR SMART; SM00282; LamG; 3.
DR SMART; SM00192; LDLa; 4.
DR SMART; SM00200; SEA; 1.
DR PROSITE; PS00022; EGF_1; 9.
DR PROSITE; PS01186; EGF_2; 6.
DR PROSITE; PS50026; EGF_3; 4.
DR PROSITE; PS50835; IG_LIKE; 22.
DR PROSITE; PS50025; LAM_G_DOMAIN; 3.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 11.
DR PROSITE; PS01209; LDLRA_1; 4.
DR PROSITE; PS50068; LDLRA_2; 4.
DR PROSITE; PS50024; SEA; 1.
KW Signal; Basement membrane; Proteoglycan; Repeat; Glycoprotein;
KW Heparan sulfate; Laminin EGF-like domain; Immunoglobulin domain;
KW Extracellular matrix; EGF-like domain; Disease mutation.
FT SIGNAL 1 21
FT CHAIN 22 4391
FT DOMAIN 80 194
FT DOMAIN 198 235
FT DOMAIN 284 320
FT DOMAIN 324 360
FT DOMAIN 367 404
FT DOMAIN 405 504
FT DOMAIN 521 530
FT DOMAIN 531 730
FT DOMAIN 731 763
FT DOMAIN 764 813
FT DOMAIN 814 871
FT DOMAIN 879 923
FT DOMAIN 924 933
FT DOMAIN 934 1125
FT DOMAIN 1126 1158
FT DOMAIN 1159 1208
FT DOMAIN 1209 1265
FT DOMAIN 1275 1324
FT DOMAIN 1325 1334
FT DOMAIN 1335 1529
FT DOMAIN 1530 1562
FT DOMAIN 1563 1612
FT DOMAIN 1613 1670
FT DOMAIN 1677 1771
FT DOMAIN 1772 1865
FT DOMAIN 1866 1955
FT DOMAIN 1956 2051
FT DOMAIN 2052 2151

FT DOMAIN 2152 2244 IG-LIKE C2-TYPE 7.
FT DOMAIN 2245 2340 IG-LIKE C2-TYPE 8.
FT DOMAIN 2341 2436 IG-LIKE C2-TYPE 9.
FT DOMAIN 2437 2533 IG-LIKE C2-TYPE 10.
FT DOMAIN 2534 2629 IG-LIKE C2-TYPE 11.
FT DOMAIN 2630 2726 IG-LIKE C2-TYPE 12.
Query Match 8.7%; Score 113; DB 1; Length 4391;
Best Local Similarity 26.5%; Pred. No. 12;
Matches 75; Conservative 27; Mismatches 83; Indels 98; Gaps 16;
QY 2 GKYNKRCRGAAGAEVAAVEVT-----QVVG--VETRSAAATGG-----VA 41
Db 2499 GEYVCRVVGSSGTQEAASLVLTQQRLSGSHSQGVAYPVRIESSASLANGHHTLDNLCLVA 2558
QY 42 KVAPR-----RKRAP-----AGEPAAAVSAG----- 62
Db 2559 SQAPHTITWYKRGSLPSRHQIVGSRRLRIPQVTPADSGEYVCHVSNAGSRETSLIVTIQ 2618
QY 63 GDGSCYIHLRSRMLFMAPPQPQPSVDSVPTPVEA-----ADGAGGQGAALAAGLSRCSS 118
Db 2619 GSGSS---HVPS---VSPPIRIES---SSPTVVEGQTLDLNCVVARQPQAIIITWYKRGGS 2669
QY 119 TAS-----SVNLGLG---GQSGHTCR---SYDAABAGDHLVDVSAASNSGSGPDRE 166
Db 2670 LPSRHQTHGSHLRLHQMVSADSGEYVCRANNIDALEAS---IVISVSPSAGSPSAPGSS 2726
QY 167 ---RRETPSSRAHGLSDLESLAG-----HKTGPSLPA 198
Db 2727 MPIRIESSSHVAEGETLDLNCVWPQGAHAQVTHWKRGGSLPS 2769

RESULT 3

IE18_PVKKA STANDARD; PRT; 1446 AA.
AC P33479;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Immediate-early protein IE180.
GN IE.
OS Pseudorabies virus (strain Kaplan) (PRV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=33703;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91021039; PubMed=2171211;
RA Vlcek C., Kozmik Z., Paces V., Schirm S., Schwyzer M.;
RT "Pseudorabies virus immediate-early gene overlaps with an oppositely
RT oriented open reading frame: characterization of their promoter and
RT enhancer regions";
RL Virology 179:365-377(1990).
CC -!- FUNCTION: THIS IE PROTEIN IS A MULTIFUNCTIONAL PROTEIN CAPABLE
CC OF MIGRATING TO THE NUCLEUS, BINDING TO DNA, TRANS-ACTIVATING
CC OTHER VIRAL GENES, AND AUTOREGULATING ITS OWN SYNTHESIS.
CC -!- SUBCELLULAR LOCATION: Nucleus of infected cells.
CC -!- PTM: A long stretch of serine residues may be a major site of
CC phosphorylation.
CC -!- SIMILARITY: BELONGS TO THE ICP4/IE140/IE180 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M34651; AAA47470.1; -.
DR PIR; A45344; A45344.
DR InterPro; IPR005205; Herpes_ICP4_C.
DR InterPro; IPR005206; Herpes_ICP4_N.

DR Pfam; PF03585; Herpes_ICP4_C; 1.
DR Pfam; PF03584; Herpes_ICP4_N; 1.
KW Early protein; Transcription regulation; Trans-acting factor;
KW DNA-binding; Phosphorylation; Nuclear protein.
FT DOMAIN 347 354 POLY-SER.
FT DOMAIN 379 397 POLY-SER.
SQ SEQUENCE 1446 AA; 148640 MW; 81F43A3DE3DDA068 CRC64;

Query Match
Best Local Similarity 8.6%; Score 112; DB 1; Length 1446;
Matches 64; Conservative 19; Mismatches 78; Indels 78; Gaps 11;

QY 9 RGAAGAEVAAVEVT-----QVGVTRSR-----AAATGGVAKVAPRRKRAPAGEPAAA 58
Db 745 RQQAADSVLVARTVAPLVRYSDGARAREAAWTYAAALFAPANVAAARLAEAAARPGPA 804

QY 59 VSAGGSGSCYIHLRSMLFMAPPQPQPSVDSVPTVEAADGA-AGQGAALAAAGLSRCS 117
Db 805 EPAPG-----LPPLWPEQGLVVPAPAPAAAGAPSGLPSPSPASTKS 849

QY 118 S-----TASSVNLGLGGQSGSHTCRSYDAAEAGDHLVDVSAANSVSGSPD-----RERR 168
Db 850 SSSTKSSSTKSLGSSG-----YASSPAAGDPAPERRKKK 887

QY 169 ETTPSSRAHGLSD-----LESFLAGHK-----TGP-----SL-----PAATPAELI 206
Db 888 RRAPGARRPGDEEGELSGAALRGDGHGRDDEEDRPRKRRLGLGPADPPAPALL 946

RESULT 4
TBC2_CHLRE
ID TBC2_CHLRE STANDARD; PRT; 1115 AA.
AC Q8VXP3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tbc2 translation factor, chloroplast precursor.
GN TBC2.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=137C / CC-125;
RX MEDLINE=22053252; PubMed=12045185;
RA Auchincloss A.H., Zerges W., Perron K., Girard-Bascou J.,
RA Rochaix J.-D.;
RT "Characterization of Tbc2, a nucleus-encoded factor specifically
required for translation of the chloroplast psbC mRNA in
Chlamydomonas reinhardtii";
RL J. Cell Biol. 157:953-962(2002).
CC -!- FUNCTION: Required for expression of the chloroplast encoded psbC
mRNA, most likely for translation initiation. Interacts with the
5'UTR of psbC.
CC -!- SUBUNIT: Part of a 400 kDa complex which is not stably associated
with RNA.
CC -!- SUBCELLULAR LOCATION: Chloroplast stroma.
CC -!- SIMILARITY: IN THE CENTRAL SECTION, TO CRP1 OF ZEA MAYS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ427966; CAD20887.1; -
KW Chloroplast; Transit peptide; Repeat.
FT TRANSIT 1 ? CHLOROPLAST (POTENTIAL).
FT CHAIN ? 1115 TBC2 TRANSLATION FACTOR.
FT DOMAIN 483 1068 9 X 38 AA APPROXIMATE REPEATS.

FT REPEAT 483 521 1.
FT REPEAT 607 645 2.
FT REPEAT 685 723 3.
FT REPEAT 724 763 4.
FT REPEAT 764 803 5.
FT REPEAT 804 842 6.
FT REPEAT 843 880 7.
FT REPEAT 990 1029 8.
FT REPEAT 1030 1068 9.
FT DOMAIN 187 210 SER-RICH.
FT DOMAIN 253 256 POLY-ASP.
FT DOMAIN 271 326 ALA-RICH.
FT DOMAIN 370 377 POLY-GLN.
FT DOMAIN 883 1115 ALA-RICH.
SQ SEQUENCE 1115 AA; 114823 MW; 2239799E91C5D8F7 CRC64;

Query Match
Best Local Similarity 8.6%; Score 111.5; DB 1; Length 1115;
Matches 70; Conservative 28; Mismatches 104; Indels 53; Gaps 11;

QY 30 RSRSAATGGVAK---VAPRRKRAPAGEPAAAVSAGDGGSCY-IHLRSMLFM----- 79
Db 164 RAGSGASTGRARGWGSGPGRN-----GSGSSSVSVNGSGSSSSSSSLAMQMQLSM 219

QY 80 -----APPQPQPSVDS-----VPTPVEAADGAAGQ-----GAALAAAGLSRCS 118
Db 220 ASIGDDVVSGVNAVGPVPSGGADALLDLEMSSILDDDDGAGARQLQMSDDLAAGLEAAAT 279

QY 119 TASSVNLGLGGQSGSHTCRSYDAAEAGDHLV-----DVSAANSVSGSPDRRERTTTPS 173
Db 280 TAAPEAGVAAAGGTGAGAAADAASSAPSLVAAAAAASPASSPDVARTLTLL 339

QY 174 SRANGELSDLESLAGHKTGPSLPAA-TPAAELIVPPAHEIQEFFFFAAEAQAQKRFASKY 232
Db 340 SRAFS--LGLDS-----LSGPQLAAVFTGLAVLRPRQQQQQQQQAAGAGANAGAGG- 391

QY 233 NFDFVRGVPLDAGGR 247
Db 392 ----VGGVGVSAAGR 402

RESULT 5
HLES_DROME
ID HLES_DROME STANDARD; PRT; 1077 AA.
AC Q02308; Q9VDK0; Q9VDK1;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Hairless protein.
GN H OR CG5460.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=92387549; PubMed=1516831;
RA Bang A.G., Posakony J.W.;
RT "The Drosophila gene Hairless encodes a novel basic protein that
controls alternative cell fates in adult sensory organ development.";
RL Genes Dev. 6:1752-1769(1992).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=93041287; PubMed=1419850;
RA Maier D., Stumm G., Kuhn K., Preiss A.;
RT "Hairless, a Drosophila gene involved in neural development, encodes
a novel, serine rich protein.";
RL Mech. Dev. 38:143-156(1992).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;

CC -!- FUNCTION: Nematode cuticles are composed largely of collagen-like
CC proteins. The cuticle functions both as an exoskeleton and as a
CC barrier to protect the worm from its environment (By similarity).
CC -!- SUBUNIT: Collagen polypeptide chains are complexed within the
CC cuticle by disulfide bonds and other types of covalent cross-links
CC (By similarity).
CC -!- SIMILARITY: Belongs to the cuticular collagen family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L11247; AAA28008.1; -.
CC PIR; S44796; S44796.
CC WormPep; F09G8.6; CE00141.
CC InterPro; IPR002486; Col_cuticle_N.
CC InterPro; IPR008160; Collagen.
CC Pfam; PF01484; Col_cuticle_N; 1.
CC Pfam; PF01391; Collagen; 3.
KW Hypothetical protein; Cuticle; Connective tissue; Repeat;
KW Multigene family; Collagen.
FT DOMAIN 91 120 TRIPLE-HELICAL REGION.
FT DOMAIN 137 199 TRIPLE-HELICAL REGION.
FT DOMAIN 202 264 TRIPLE-HELICAL REGION.
SQ SEQUENCE 278 AA; 26228 MW; E21A3B3EB925FA27 CRC64;

Query Match 8.2%; Score 107; DB 1; Length 278;

Best Local Similarity 25.5%; Pred. No. 1.6;
Matches 47; Conservative 11; Mismatches 60; Indels 66; Gaps 6;

QY 10 GAAGAEEVAEEVTQVVGVRTRSRSAATGGVAKVAPRRKRAPAGEPAAVAVSAGDGGSCY 69
Db 112 GAAGIDGVAIQF-----ASAAGGACIQCPAGEAGPAGAPGAPGPDG----- 155

QY 70 IHLRSMLFMAPPQPPSVD---SVPTPVEAADGAAGQQAALAAGLSRCSSTASSVNLG 126
Db 156 -----QPGADGGGAGPAPA-GPEGPAGDAGAPGAPG-----APGNDG 191

QY 127 LGQGRSHTCRSYDAEAGGDH-----VLVDVSAASNSGSGP 163
Db 192 QPGQNGQRSTGTGPAAGAPGPQGPVGSQPGSAGAPGAPGAPGVDGPGANGPQP 251

QY 164 DRER 167

Db 252 DGEQ 255

RESULT 7

HUTG_HALN1 STANDARD; PRT; 306 AA.
AC Q9HQD7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Probable formiminoglutamate (EC 3.5.3.8) (Formiminoglutamate)
DE (Formiminoglutamate hydrolase).
GN HUTG OR VNG1209G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,

RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -!- CATALYTIC ACTIVITY: N-formimidoyl-L-glutamate + H(2)O = L-
CC glutamate + formamide.
CC -!- COFACTOR: Binds 2 manganese ions per subunit (By similarity).
CC -!- PATHWAY: Histidine degradation; fourth step.
CC -!- SIMILARITY: Belongs to the arginase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB005047; AAG19578.1; -.
CC PIR; F84276; F84276.
CC HAMAP; MF_00737; -; 1.
CC InterPro; IPR006035; Arg_agm_form.
CC InterPro; IPR005923; HutG.
CC Pfam; PF00491; arginase; 1.
CC TIGRfams; TIGR01227; hutG; 1.
CC PROSITE; PS00147; ARGINASE_1; FALSE_NEG.
CC PROSITE; PS00148; ARGINASE_2; 1.
CC PROSITE; PS01053; ARGINASE_3; FALSE_NEG.
KW Hydrolase; Histidine metabolism; Metal-binding; Manganese;
KW Complete proteome.

FT METAL 123 123 MANGANESE 1 (BY SIMILARITY).
FT METAL 147 147 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 149 149 MANGANESE 2 (BY SIMILARITY).
FT METAL 151 151 MANGANESE 1 (BY SIMILARITY).
FT METAL 234 234 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 236 236 MANGANESE 2 (BY SIMILARITY).
SQ SEQUENCE 306 AA; 31422 MW; FD98BD7DBEDF958 CRC64;

Query Match 8.1%; Score 106; DB 1; Length 306;

Best Local Similarity 25.8%; Pred. No. 2.1;

Matches 70; Conservative 20; Mismatches 115; Indels 66; Gaps 13;

QY 9 RGAAGAEVAEEVTQVVGVRTRSRSAATGGVA-----KVAPRRKRAPAG 53

Db 53 RGARDGPSAIRE--SLAGVKTHFDAGAVSGVADLGDVLPDGDVADTQAAVREAAAEVH 110

QY 54 EPAA-AVSAGDGGSCYIHLRSMLFMAPPQPPSVDSPVTPVEADG-----AAGQGA 107

Db 111 ETAALFVFGDNSLSYANV-----APLVAADNGAVGVSVDAHLDCRAVGDGRGP 160

QY 108 A-----LAAGLSRCSSTASSVNLGLGQGRSHTCRSYDA--AEAGGDHVLVDVSAAS 157

Db 161 TSGTPYRQLFDAGLDA-----LAVVGARHPETTTTAYAGFLRDQGGRIVTSDAVAAD 211

QY 158 NSGSGPDRERRETTTPSSRAHAGELSDLESDLAGHKTPSLPAATPAAELIVPPAHEIQEFF 217

Db 212 RDGS-LDAAREALDGVHDVVSV-DID-----VLDAAYPGASAPTPGGIQPRELF 259

QY 218 AAAEAAQAKRFASKYKYNFDFVRGVP-LDAGGR 247

Db 260 ALVEALAAASDDRIR-GPELVETAPTLDTGGR 289

RESULT 8

OGFR_HUMAN STANDARD; PRT; 677 AA.
ID OGFR_HUMAN
AC Q9NZT2; Q96029; Q96CM2; Q9BQW1; Q9H4H0; Q9H7J5; Q9NZT3; Q9NZT4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Opioid growth factor receptor (OGFr) (Zeta-type opioid receptor) (7-60
DE protein).

GN OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Placenta;
RX MEDLINE=20143540; PubMed=10677613;
RA Zagon I.S., Verderame M.F., Allen S.S., McLaughlin P.J.;
RT "Cloning, sequencing, chromosomal location, and function of cDNAs
RT encoding an Opioid growth factor receptor (OGFR) in humans.";
RL Brain Res. 856:75-83 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Takanosu M., Liu J., Mayne R., Wood B.M., Brewton R.G.;
RT "Genomic structure of human gene 7-60.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark A.G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Copley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leharasliho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConnachie L.J., McLay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871 (2001).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny K.D., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [5]
RP SEQUENCE OF 37-677 FROM N.A. (ISOFORM 1).
RC TISSUE=Spleen;
RX MEDLINE=21082933; PubMed=11214971;
RA Hattori A., Okumura K., Nagase T., Kikuno R., Hirose M., Ohara O.;
RT "Characterization of long cDNA clones from human adult spleen.";
RL DNA Res. 7:357-366 (2000).
RN [6]
RP REVIEW.
RX MEDLINE=21888341; PubMed=11890982;
RA Zagon I.S., Verderame M.F., McLaughlin P.J.;
RT "The biology of the opioid growth factor receptor (OGFR).";
RL Brain Res. Brain Res. Rev. 38:351-376 (2002).
CC -!- FUNCTION: Receptor for opioid growth factor (OGF), also known as
CC Met-enkephalin. Seems to be involved in growth regulation.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity). The OGF/OGFR
CC complex is probably translocated to the nucleus.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9NZT2-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9NZT2-2; Sequence=VSP_004060;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Highly expressed in the heart and liver,
CC moderately in skeletal muscle and kidney and to a lesser extent in
CC brain and pancreas. Expressed in fetal tissues including liver and
CC kidney.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to frameshifts
CC in position 1 and 181.
CC -!- CAUTION: Ref.3 (CAC28882) sequence differs from that shown due to
CC erroneous gene model prediction.
CC -!- CAUTION: Ref.3 (CAC12749) sequence differs from that shown due to
CC a frameshift in position 514.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF172451; AAF64404.1; -;
DR EMBL; AF172452; AAF64405.1; -;
DR EMBL; AF172453; AAF64406.1; -;
DR EMBL; AF112980; AAD03745.1; ALT_FRAME.
DR EMBL; AF109134; AAD03737.1; ALT_FRAME.
DR EMBL; AL035669; CAC28882.1; ALT_SEQ.
DR EMBL; AL035669; CAC12749.1; ALT_FRAME.
DR EMBL; BC014137; AAH14137.1; -;
DR EMBL; AK024485; BAB15775.1; -;
DR Genew; HGNC:15768; OGFR.
DR MIM; 606459; -;
DR GO; GO:0004985; F:opioid receptor activity; NAS.
DR GO; GO:0001558; P:regulation of cell growth; NAS.
DR InterPro; IPR006757; OGFR_N.
DR InterPro; IPR006770; OGFR_repeat.
DR Pfam; PF04664; OGFR_N; 1.
DR Receptor; Growth regulation; Repeat; Alternative splicing.
FT DOMAIN 267 283 BIPARTITE NUCLEAR LOCALIZATION SIGNAL
FT (POTENTIAL).
FT 7 X APPROXIMATE 20 AA TANDEM REPEATS.
FT DOMAIN 517 656
FT REPEAT 517 536 1.
FT REPEAT 537 556 2.
FT REPEAT 557 576 3.
FT REPEAT 577 596 4.
FT REPEAT 597 616 5.
FT REPEAT 617 636 6.
FT REPEAT 637 656 7.
FT VARSPLIC 558 577 Missing (in isoform 2).
FT /FTId=VSP_004060.
FT CONFLICT 27 28 DG -> EA (IN REF. 2).

RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA Sampei G., Seki Y., Sivasundaram S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.,
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
CC -!- SIMILARITY: Belongs to the tail fiber family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000234; AAC74454.1; ALT_INIT.
DR EMBL; D90774; BAA14966.1; -.
DR EMBL; D90775; BAA14975.1; -.
DR PIR; G64887; G64887.
DR EcoGene; EGI3370; stfR.
DR InterPro; IPR008969; CarboxypepD_reg.
DR InterPro; IPR005003; Phage_fiber_1.
DR InterPro; IPR005068; Phage_fiber_2.
DR Pfam; PF03335; Phage_fiber; 6.
DR Pfam; PF03406; Phage_fiber_2; 1.
KW Hypothetical protein; Fiber protein; Repeat; Complete proteome.
SQ SEQUENCE 1120 AA; 113779 MW; 542E59D71BE795B4 CRC64;

Query Match 8.1%; Score 105.5; DB 1; Length 1120;
Best Local Similarity 22.4%; Pred. No. 8.9;
Matches 56; Conservative 30; Mismatches 113; Indels 51; Gaps 6;

QY 9 RGAAGAEVAA--VEVTQVGVTRSRSAATGGVAKVAPRRK-RAPAGEPAAAVSAGDGG 65
Db 237 RDAASKEAAKSSSETNASSASSATAAGNSAKAAKTSETNARSSETAAGQSASAAA 296
QY 66 GSCYIHLRSRMLFMAPPQPPQPSVDSVPTPVEAADGAAGQQAALAGLSRCSSTASSVN- 124
Db 297 G-----SKTAAASSASAASSTAGQASASATAGKSAEASAST 336
QY 125 ----LGLGGQRGHTRCSYDAE-----AGGDHVLVDVSAASNSGSG 162
Db 337 ATTGAGEATEQASAAARSASAAKTSETNKASETSAESSKTAASASSASSASSASAS 396
QY 163 PDRERRETPSSRAHGELESLDLGHKGTG-----PSLPAATPAELIIVPAHEIQEFF 217
Db 397 KDEATRQASAAKSSATTASTKATEAAGSATAAAGSKSTAESAATRAETAAKRAEDIASAV 456
QY 218 AAEEAAQAKR 227
Db 457 ALEDASTTKK 466

RESULT 11
FXC1_MOUSE
ID FXC1_MOUSE STANDARD; PRT; 553 AA.
AC Q61572; O88409; Q61582;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Forkhead box protein C1 (Forkhead-related protein FKHL7) (Forkhead-
DE related transcription factor 3) (FREAC-3) (Transcription factor FKHL-1)
DE (Mesoderm/mesenchyme forkhead 1) (MF-1).
GN FOXC1 OR FKHL7 OR FREAC3 OR FKHL1 OR MF1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=98297351; PubMed=9635428;
RA Kume T., Deng K.Y., Winfrey V., Gould D.B., Walter M.A., Hogan B.L.M.;
RT "The forkhead/winged helix gene Mf1 is disrupted in the pleiotropic
RT mouse mutation congenital hydrocephalus.";
RL Cell 93:985-996(1998).
RN [2]
RP SEQUENCE OF 69-179 FROM N.A.
RC STRAIN=129;
RX MEDLINE=93361500; PubMed=7689224;
RA Kaestner K.H., Lee K.H., Schloendorff J., Hiemisch H.,
RA Monaghan A.P., Schuetz G.;
RT "Six members of the mouse forkhead gene family are developmentally
RT regulated.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:7628-7631(1993).
RN [3]
RP SEQUENCE OF 71-187 FROM N.A.
RX MEDLINE=93387221; PubMed=8375339;
RA Sasaki H., Hogan B.L.;
RT "Differential expression of multiple fork head related genes during
RT gastrulation and axial pattern formation in the mouse embryo.";
RL Development 118:47-59(1993).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Expressed in many embryonic tissues, including
CC prechondrogenic mesenchyme, pericardial mesenchyme, meninges,
CC endothelial cells, and kidney.
CC -!- DEVELOPMENTAL STAGE: Expressed during embryogenesis.
CC -!- SIMILARITY: Contains 1 fork-head domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF045017; AAC24209.1; -.
DR EMBL; L10406; AAA03159.1; -.
DR EMBL; X71939; CAA50741.1; -.
DR PIR; I49674; I49674.
DR HSSP; Q63245; 2HFH.
DR TRANSFAC; T02426; -.
DR MGD; MGI:1347466; Foxc1.
DR InterPro; IPR001766; TF_Fork_head.
DR Pfam; PF00250; Fork_head; 1.
DR PRINTS; PR00053; FORKHEAD.
DR ProDom; PD000425; TF_Fork_head; 1.
DR SMART; SM00339; FH; 1.
DR PROSITE; PS00657; FORK HEAD 1; 1.
DR PROSITE; PS00658; FORK HEAD 2; 1.
DR PROSITE; PS00039; FORK HEAD 3; 1.
KW DNA-binding; Nuclear protein; Transcription regulation.
FT DOMAIN 28 33 POLY-ALA.
FT DNA_BIND 77 168 FORK-HEAD.
FT DOMAIN 169 173 POLY-ARG.
FT DOMAIN 194 197 POLY-PRO.
FT DOMAIN 264 274 POLY-SER.
FT DOMAIN 375 386 POLY-GLY.
FT DOMAIN 444 451 POLY-SER.
FT DOMAIN 453 456 POLY-GLY.
FT DOMAIN 486 496 POLY-ALA.
FT CONFLICT 180 186 VKDKEEK -> KKEITFI (IN REF. 3).
SQ SEQUENCE 553 AA; 56953 MW; 3CDD12F69CA4F217 CRC64;

Query Match 8.1%; Score 105; DB 1; Length 553;
Best Local Similarity 22.9%; Pred. No. 4.6;
Matches 63; Conservative 33; Mismatches 103; Indels 76; Gaps 13;

QY 31 SRSAATGGVAKVAPRRKRAPAGEPAAAVSAGDGGSCYIHLRSRMLFMA----PPQPQP 86
Db 243 SPAAALGSGSAATVPKIESPDSS--SSSLSSGSSPPGSLPSARPLSLDAAEPPPPQAP 300
QY 87 -----SVDSVPTPVEAA-DGAAGQQAALAGLSRCSSTASSVNLGLG----GQRG- 132

Db 301 PPHSQGFSVDNMTSLRGSPQGSAAELGSLASAAASSRAGIAPPLALGAYSPQSSSL 360
 QY 133 -SHTC-RSYDAEAGDHLVDVSAASNSGSGDPRERETTPSSRAHAGELSDLE---SD 186
 Db 361 YSSPCSQSSSAGSSGG-----GGGGGGGGSSSSAAGTGAATYHCNLOAMSLYAAGE 413
 QY 187 LAGHKTGP-----SLPAATPAABLIV-----PPAHE--IQEF 216
 Db 414 RGHLLQGPAGGAGAAVDDPLPDYSLPPATSSSSSSLSHGSGGQEAASHHPASHQGRLTWS 473
 QY 217 F-----AAEAAQAQKRFASKYNFDFVR 238
 Db 474 YLNQAGDGLHLASAAAAAAYPGQQQNFHVR 508

RESULT 12
 ICP4_HSVB STANDARD; PRT; 1487 AA.
 AC P28925;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE Trans-acting transcriptional protein ICP4 (155 kDa immediate-early protein).
 DE IE OR 64.
 GN Equine herpesvirus type 1 (strain Ab4p) (EHV-1).
 OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=31520;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92295566; PubMed=1318606;
 RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
 RT "The DNA sequence of equine herpesvirus-1.";
 RL Virology 189:304-316(1992).
 CC -!- FUNCTION: THIS IE PROTEIN IS A MULTIFUNCTIONAL PROTEIN CAPABLE
 CC OF MIGRATING TO THE NUCLEUS, BINDING TO DNA, TRANS-ACTIVATING
 CC OTHER VIRAL GENES, AND AUTOREGULATING ITS OWN SYNTHESIS.
 CC -!- SUBCELLULAR LOCATION: Nucleus of infected cells.
 CC -!- PTM: A long stretch of serine residues may be a major site of
 CC phosphorylation.
 CC -!- SIMILARITY: BELONGS TO THE ICP4/IE140/IE180 FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; M86664; AAB02515.1; -.
 EMBL; M86664; AAB02499.1; -.
 PIR; A36802; EDBEF6.
 TRANSFAC; T00401; -.
 InterPro; IPR005205; Herpes_ICP4_C.
 InterPro; IPR005206; Herpes_ICP4_N.
 Pfam; PF03585; Herpes_ICP4_C; 1.
 Pfam; PF03584; Herpes_ICP4_N; 1.
 KW Early protein; Transcription regulation; Trans-acting factor;
 DNA-binding; Phosphorylation; Nuclear protein.
 FT DOMAIN 181 213 SER-RICH.
 FT DOMAIN 922 931 ARG/LYS-RICH (BASIC).
 SQ SEQUENCE 1487 AA; 154868 MW; 1D4397838D03680D CRC64;

Query Match 8.0%; Score 104.5; DB 1; Length 1487;
 Best Local Similarity 23.2%; Pred. No. 14;
 Matches 61; Conservative 23; Mismatches 88; Indels 91; Gaps 12;
 QY 26 GVRTRSRSAATGG-VAKVAPRRKRAPAGEPAAVSA----- 61
 Db 79 GKAKRSPSAGSGGPPFAAAQAPASPAPGLAAMLKMHSSVAPGNGRRATGSSSP 138

QY 62 -GGDGG-----SCYIHLRSLFMAPPQPSVDSVPTVEAADGAAGQQAALA 110
 Db 139 GGGDAADPVALDSDTETC-----PGSPQPEF-----PSSASPGGSGSPAPRVRS 181
 QY 111 AGLSRCSSSTASSVNL-----GLGGQSGSHTCSVDAAEAGGD-----HVLVDVSA 155
 Db 182 ISISSSSSSSSMDEDDQADGAGASSSSSSSSSDSDSDEGGEETPRPHSQNAKTPSA 241
 QY 156 ASNSG--SGPDRERR-ETTPSSRAHAGELSDLESLAGHKTGTPSLPAATPAEL----- 205
 Db 242 AGSPGSSGGDRPAAGAAATPKSCRSGAAS-----PGAPAPAPASAPAPSRPGG 289
 QY 206 -IVPPAHEIQEFF-AAEAAQAQK 226
 Db 290 GLLPPGARILEYLEGVREANLAK 312

RESULT 13
 ICP4_HSVK STANDARD; PRT; 1487 AA.
 AC P17473;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE Trans-acting transcriptional protein ICP4 (155 kDa immediate-early protein).
 DE IE OR 64.
 GN Equine herpesvirus type 1 (strain Kentucky A) (EHV-1).
 OS Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89370304; PubMed=2549711;
 RA Grundy F.J., Baumann R.P., O'Callaghan D.J.;
 RT "DNA sequence and comparative analyses of the equine herpesvirus type
 RT 1 immediate early gene.";
 RL Virology 172:223-236(1989).
 RN [2]
 RP SEQUENCE OF 1432-1487 FROM N.A.
 RX MEDLINE=90064773; PubMed=2555546;
 RA Harty R.N., Colle C.F. III, Grundy F.J., O'Callaghan D.J.;
 RT "Mapping the termini and intron of the spliced immediate-early
 RT transcript of equine herpesvirus 1.";
 RL J. Virol. 63:5101-5110(1989).
 CC -!- FUNCTION: THIS IE PROTEIN IS A MULTIFUNCTIONAL PROTEIN CAPABLE
 CC OF MIGRATING TO THE NUCLEUS, BINDING TO DNA, TRANS-ACTIVATING
 CC OTHER VIRAL GENES, AND AUTOREGULATING ITS OWN SYNTHESIS.
 CC -!- SUBCELLULAR LOCATION: Nucleus of infected cells.
 CC -!- PTM: A long stretch of serine residues may be a major site of
 CC phosphorylation.
 CC -!- SIMILARITY: BELONGS TO THE ICP4/IE140/IE180 FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; J04366; AAA46089.1; -.
 EMBL; M30498; AAA66554.1; -.
 PIR; A33764; EDBEEL.
 InterPro; IPR005205; Herpes_ICP4_C.
 InterPro; IPR005206; Herpes_ICP4_N.
 Pfam; PF03585; Herpes_ICP4_C; 1.
 Pfam; PF03584; Herpes_ICP4_N; 1.
 KW Early protein; Transcription regulation; Trans-acting factor;
 DNA-binding; Phosphorylation; Nuclear protein.
 FT DOMAIN 181 213 SER-RICH.
 FT DOMAIN 922 931 ARG/LYS-RICH (BASIC).


```
SQ SEQUENCE 1487 AA; 154716 MW; 044E39A570608A6B CRC64;
Query Match 8.0%; Score 104.5; DB 1; Length 1487;
Best Local Similarity 23.2%; Pred. No. 14;
Matches 61; Conservative 23; Mismatches 88; Indels 91; Gaps 12;

QY 26 GVRTRSRSAATGG-VAKVAPRRKRAPAGEPAAAVSA----- 61
Db 79 GKAKRSPSAAGSGPPPTPAAQAQSPASPAPGLAAMLKMHSSVAPGNGRATGSSSP 138
QY 62 -GGDGG-----SCYIHLRSRLFMAPPQPSVDSVPTPVEAADGAGCQQAALA 110
Db 139 GGGDAADPVALDSTETC-----PGSPQPEF-----PSSASPGGSGSPAPRVRS 181
QY 111 AGLSRCSSTASSVNL-----GLGGQGSHTCRSYDAEAGGD-----HVLVDVSA 155
Db 182 ISISSSSSSSMDEDDQADGAGASSSSSSSSSDSDSDEGGEETPRPHSQNAKTPSA 241
QY 156 ASNSG--SGPDRER-ETTPSSRAHGLSDLESIDLGHKTPSLPAATPAEL----- 205
Db 242 AGSPGPGSGDRPAAAGATPKCRSGAAS-----PGAPAPAPASAPAPSRPGG 289
QY 206 -IVPPAHEIQEFF-AAAEAAQAK 226
Db 290 GLPPGARILEYLEGVREANLAK 312

RESULT 14
BASP_RAT
ID BASP_RAT STANDARD; PRT; 219 AA.
AC Q05175;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Brain acid soluble protein 1 (BASP1 protein) (Neuronal axonal membrane
protein NAP-22).
GN BASP1 OR NAP22.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=93293900; PubMed=8390468;
RA Maekawa S., Maekawa M., Hattori S., Nakamura S.;
RT "Purification and molecular cloning of a novel acidic calmodulin
binding protein from rat brain."
RL J. Biol. Chem. 268:13703-13709(1993).
RN [2]
RP MYRISTOYLATION.
RX MEDLINE=94250689; PubMed=8193160;
RA Maekawa S., Matsuura Y., Nakamura S.;
RT "Expression and myristoylation of NAP-22 using a baculovirus transfer
vector system."
RL Biochim. Biophys. Acta 1218:119-122(1994).
RN [3]
RP PARTIAL SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=97455984; PubMed=9310187;
RA Mosevitsky M.I., Capony J.-P., Skladchikova G.Y.U., Novitskaya V.A.,
RA Plekhanov A.Y.U., Zakharov V.V.;
RT "The BASP1 family of myristoylated proteins abundant in axonal
termini. Primary structure analysis and physico-chemical
properties."
RL Biochimie 79:373-384(1997).
CC -!- SUBCELLULAR LOCATION: MEMBRANE-ANCHORED. ASSOCIATED WITH THE
CC MEMBRANES OF "GROWTH CONES" THAT FORM THE TIPS OF ELONGATING
CC AXONS.
CC -!- TISSUE SPECIFICITY: Brain.
CC -!- SIMILARITY: BELONGS TO THE BASP/NAP22 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D14441; BAA03333.1; -.
DR PIR; A46597; A46597.
DR InterPro; IPR008408; BASP1.
DR Pfam; PF05466; BASP1; 1.
KW Membrane; Myristate; Neurone; Calmodulin-binding; Lipoprotein.
FT INIT MET 0
FT LIPID 1 N-myristoyl glycine.
SQ SEQUENCE 219 AA; 21659 MW; 7E79A441E021D163 CRC64;

Query Match 8.0%; Score 104; DB 1; Length 219;
Best Local Similarity 26.2%; Pred. No. 2;
Matches 59; Conservative 22; Mismatches 92; Indels 52; Gaps 10;

QY 6 RKCRGAAGAE-----VAAVEVTQVVGVRTRSRSAATGGV-----AKVAPRRKRA 50
Db 23 KKAEGAGTEEGTQKESEPPQAAADATEVKESAEEKPKDAADGEAKAEKEADKAAKEEA 82
QY 51 PAGEPAAAVSAGDGGSCYIHLRSRLFMAPPQPSVDSVPTPVE--AADG-AAGQQA 107
Db 83 PKAEPEKSEGA-----AEEQPEPA-----PAPEQEAAPGPAAG--GE 118
QY 108 ALAAGLSRCSSTASSVNLGLGGQGSHTCRSYDAEAGGDHVLVDVSAASNSGSGPDRER 167
Db 119 APKAG---EASAESTGAADGAPQEEGEAKTEAPAAAGPEAKSDAAPAASDS-----KPS 169
QY 168 RETTPSSRAHGLSDLESIDLGHKTPSLPAATPAELIVPPAHE 212
Db 170 TEPAPSSKETPAASEAPSS-AAKAPAPAAPAAEPQAEAPVASSEQ 213

RESULT 15
KLTK_MOUSE
ID KLTK_MOUSE STANDARD; PRT; 888 AA.
AC P08923;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Leukocyte tyrosine kinase receptor precursor (EC 2.7.1.112).
GN LTK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC STRAIN=BALB/c;
RX MEDLINE=90291994; PubMed=2357970;
RA Bernards A., de la Monte S.;
RT "The ltk receptor tyrosine kinase is expressed in pre-B lymphocytes
and cerebral neurons and uses a non-AUG translational initiator."
RL EMBO J. 9:2279-2287(1990).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC STRAIN=BALB/c;
RX MEDLINE=92115335; PubMed=1662793;
RA Haase V.H., Snijders A.J., Cooke S.M., Teng M.N., Kaul D.,
RA le Beau M.M., Bruns G.A., Bernards A.;
RT "Alternatively spliced ltk mRNA in neurons predicts a receptor with a
larger putative extracellular domain."
RL Oncogene 6:2319-2325(1991).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM A).
RX MEDLINE=88232962; PubMed=2836739;
RA Ben-Neriah Y., Bauskin A.R.;
RT "Leukocytes express a novel gene encoding a putative transmembrane
protein-kinase devoid of an extracellular domain."
RT
```


RL Nature 333:672-676(1988).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS A; B; C AND D).
RX MEDLINE=93141274; PubMed=8380920;
RA Snijders A.J., Haase V.H., Bernards A.;
RT "Four tissue-specific mouse ltk mRNAs predict tyrosine kinases that
RT differ upstream of their transmembrane segment.";
RL Oncogene 8:27-35(1993).
RN [5]
RP SUBCELLULAR LOCATION, AND INTERACTION WITH CALNEXIN (ISOFORM A).
RX MEDLINE=97150899; PubMed=8995435;
RA Snijders A.J., Ho S.C., Haase V.H., Pillai S., Bernards A.;
RT "A lymphocyte-specific ltk tyrosine kinase isoform is retained in the
RT endoplasmic reticulum in association with calnexin.";
RL J. Biol. Chem. 272:1297-1301(1997).
CC -|- FUNCTION: THE EXACT FUNCTION OF THIS PROTEIN IS NOT KNOWN, IT IS
CC PROBABLY A RECEPTOR WITH A TYROSINE-PROTEIN KINASE ACTIVITY.
CC -|- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -|- SUBUNIT: Isoform A binds calnexin.
CC -|- SUBCELLULAR LOCATION: Type I membrane protein. Isoform A and
CC isoform C are retained in the endoplasmic reticulum.
CC -|- ALTERNATIVE PRODUCTS:
CC Event=Alternative promoter;
CC Comment=2 isoforms, A and B, may be produced by use of
CC alternative promoters;
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=Additional isoforms seem to exist;
CC Name=D;
CC IsoId=P08923-1; Sequence=Displayed;
CC Name=A;
CC IsoId=P08923-2; Sequence=VSP_002950, VSP_002951, VSP_002952;
CC Note=May be produced by use of alternative promoters. Starts at
CC a CUG codon;
CC Name=B;
CC IsoId=P08923-3; Sequence=VSP_002950, VSP_002951;
CC Note=May be produced by use of alternative promoters. Starts at
CC a CUG codon;
CC Name=C;
CC IsoId=P08923-4; Sequence=VSP_002952;
CC -|- TISSUE SPECIFICITY: Subsets of lymphoid and neuronal cells.
CC -|- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin
CC receptor subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X52621; CAA36848.1; -.
DR EMBL; X07984; CAA30793.1; ALT_INIT.
DR EMBL; M90470; AAA39451.1; -.
DR PIR; I58378; I58378.
DR PIR; S12792; S12792.
DR HSSP; P11362; 1FGK.
DR MGD; MGI:96840; Ltk.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002011; RecepttyrkinII.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR008266; Tyr_pkinase_AS.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
DR PROSITE; PS00239; RECEPTOR TYR KIN II; 1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
DR Transferase; Tyrosine-protein kinase; Transmembrane; Receptor;
KW Endoplasmic reticulum; ATP-binding; Phosphorylation; Glycoprotein;

KW Signal; Alternative splicing; Alternative promoter usage.
FT SIGNAL 1 16
FT CHAIN 17 888
FT DOMAIN 17 421
FT TRANSMEM 422 446
FT DOMAIN 447 588
FT DOMAIN 506 782
FT NP BIND 512 520
FT BINDING 540 540
FT ACT SITE 639 639
FT MOD_RES 672 672
FT CARBOHYD 377 377
FT CARBOHYD 409 409
FT VARSPLIC 1 251
FT VARSPLIC 252 252
FT VARSPLIC 271 331
FT CONFLICT 789 789
FT CONFLICT 875 875
FT SEQUENCE 888 AA; 94436 MW; 3FFCA80AB4863C55 CRC64;
SQ
Query Match 8.0%; Score 104; DB 1; Length 888;
Best Local Similarity 28.0%; Pred.No. 8.8;
Matches 60; Conservative 16; Mismatches 94; Indels 44; Gaps 11;
QY 8 CRGAAGAEVAAEVTVGVTRSRSAATGGVAKVAPRRKRAPAGEPAAAVSAGDGGG 67
Db 179 CLGESGEHATTYGTERTIPGWRWAGGGGGGATSIIF---RLRAGEPEPLLVAAAGGGGR 234
QY 68 CYIHLRSRMLFMAPPQPQPSVDSVPTPVE---AADGAAGQQGAALAAAGLSRCSSTASSVN 124
Db 235 SY---RRR-----PDRGRTOAVPERLETRAAAPGSGGRGGA--AGGSGGWTSTRAHSPQ 282
QY 125 LGLG---GQSGSHTCRSYDA---AEAGGDHVLVDVSAASNSGSGGPDRETRTPSSRAH 177
Db 283 AGRSPPREGAEGGEGCAEAWAALRWAAAGGFGGGGACAAAGGGGGG-----YRG 330
QY 178 GELSDLES DL--AGHKTGPSLPAATPAEELIVPP 209
Db 331 GDT S--ESDLLWADGEDGT SF--VHPSGELYLQP 360
Search completed: October 1, 2004, 16:51:44
Job time : 26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 16:45:36 ; Search time 116 Seconds
(without alignments)
696.316 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 1304
Sequence: 1 MGKYMRCRGAAGAEVAAVE.....VRGVPLDAGGRFEWAPVVS I 256

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL, 25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 225 | 17.3 | 242 | 10 Q9FW65 | Q9fw65 oryza sativ |
| 2 | 225 | 17.3 | 242 | 10 Q7XDH8 | Q7xdh8 oryza sativ |
| 3 | 212.5 | 16.3 | 222 | 10 Q9FKB5 | Q9fkb5 arabidopsis |
| 4 | 207 | 15.9 | 185 | 10 Q8GT28 | Q8gt28 lycopersico |
| 5 | 179.5 | 13.8 | 286 | 10 Q94CM0 | Q94cm0 arabidopsis |
| 6 | 178 | 13.7 | 196 | 10 Q48597 | Q48597 chenopodium |
| 7 | 178 | 13.7 | 289 | 10 Q48846 | Q48846 arabidopsis |
| 8 | 177.5 | 13.6 | 286 | 10 Q8GYJ3 | Q8gyj3 arabidopsis |
| 9 | 170 | 13.0 | 192 | 10 Q9FS28 | Q9fs28 pisum sativ |
| 10 | 169.5 | 13.0 | 210 | 10 Q8GT29 | Q8gt29 lycopersico |
| 11 | 163 | 12.5 | 207 | 10 Q8GUA2 | Q8gua2 nicotiana t |
| 12 | 153 | 11.7 | 189 | 10 Q9LRY0 | Q9lry0 arabidopsis |
| 13 | 152 | 11.7 | 163 | 10 Q93V92 | Q93v92 nicotiana t |
| 14 | 134 | 10.3 | 196 | 10 Q9LJL5 | Q9lj15 arabidopsis |
| 15 | 129.5 | 9.9 | 204 | 10 Q8S9U7 | Q8s9u7 oryza sativ |
| 16 | 129.5 | 9.9 | 1156 | 16 Q9Z5A4 | Q9z5a4 streptomyce |

| | | | | | |
|----|-------|-----|------|-----------|--------------------|
| 17 | 128 | 9.8 | 801 | 5 Q23635 | Q23635 caenorhabdi |
| 18 | 126 | 9.7 | 1080 | 10 Q9SDC1 | Q9sdc1 oryza sativ |
| 19 | 124 | 9.5 | 826 | 12 P89473 | P89473 herpes simp |
| 20 | 121.5 | 9.3 | 195 | 10 Q9FX90 | Q9fx90 arabidopsis |
| 21 | 121.5 | 9.3 | 195 | 10 Q94CL9 | Q94cl9 arabidopsis |
| 22 | 121.5 | 9.3 | 1151 | 13 Q57580 | Q57580 gallus gall |
| 23 | 121.5 | 9.3 | 1340 | 16 Q9L1H8 | Q9llh8 streptomyce |
| 24 | 121.5 | 9.3 | 2639 | 5 Q76786 | Q76786 antheraea p |
| 25 | 121 | 9.3 | 191 | 10 Q82809 | Q82809 arabidopsis |
| 26 | 121 | 9.3 | 191 | 10 Q8LDX1 | Q8ldx1 arabidopsis |
| 27 | 121 | 9.3 | 191 | 10 Q04154 | Q04154 arabidopsis |
| 28 | 121 | 9.3 | 2936 | 5 Q9NKP7 | Q9nkp7 leishmania |
| 29 | 119 | 9.1 | 1787 | 10 Q9M4X9 | Q9m4x9 chlamydomon |
| 30 | 118.5 | 9.1 | 775 | 16 Q9F342 | Q9f342 streptomyce |
| 31 | 117.5 | 9.0 | 1468 | 5 Q9GUB5 | Q9gub5 galleria me |
| 32 | 116.5 | 8.9 | 452 | 16 Q9KY45 | Q9ky45 streptomyce |
| 33 | 116.5 | 8.9 | 605 | 2 Q9AH41 | Q9ah41 neisseria c |
| 34 | 116.5 | 8.9 | 1387 | 16 Q82C89 | Q82c89 streptomyce |
| 35 | 116.5 | 8.9 | 2655 | 5 Q964F4 | Q964f4 antheraea y |
| 36 | 116 | 8.9 | 362 | 10 Q8S031 | Q8s031 oryza sativ |
| 37 | 115.5 | 8.9 | 1338 | 4 Q9Y6I5 | Q9y6i5 homo sapien |
| 38 | 115 | 8.8 | 156 | 10 Q93YF6 | Q93yf6 nicotiana s |
| 39 | 115 | 8.8 | 1414 | 5 Q26634 | Q26634 strongyloce |
| 40 | 114.5 | 8.8 | 812 | 16 Q8ZQ81 | Q8zq81 salmonella |
| 41 | 114.5 | 8.8 | 832 | 16 Q9KZY7 | Q9kzy7 streptomyce |
| 42 | 113.5 | 8.7 | 467 | 16 Q06256 | Q06256 mycobacteri |
| 43 | 113.5 | 8.7 | 467 | 16 Q7TWI1 | Q7twi1 mycobacteri |
| 44 | 113 | 8.7 | 1541 | 5 Q15837 | Q15837 leishmania |
| 45 | 112.5 | 8.6 | 3012 | 5 Q97205 | Q97205 leishmania |

ALIGNMENTS

RESULT 1

Q9FW65 ID Q9FW65 PRELIMINARY; PRT; 242 AA.
AC Q9FW65;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative cyclin-dependent kinase inhibitor.
GN OSJNB0094K03.16.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,
RA Hsiao J., Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,
RA Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,
RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 10 BAC OSJNB0094K03 genomic sequence.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC069145; AAG16867.1; -.
DR Gramene; Q9FW65; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004861; F:cyclin-dependent protein kinase inhibitor a. . .; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0007050; P:cell cycle arrest; IEA.
DR InterPro; IPR003175; CDI.
DR Pfam; PF02234; CDI; 1.
KW Cyclin; Kinase.
SQ SEQUENCE 242 AA; 26750 MW; F215A01234735928 CRC64;

Query Match 17.3%; Score 225; DB 10; Length 242;
Best Local Similarity 28.0%; Pred. No. 5.9e-08;
Matches 84; Conservative 24; Mismatches 80; Indels 112; Gaps 12;

QY 1 MGKYMRCRGAAGAEVAAVEVTQV----VGVRTRSRSAATGVAKVAPRRKRAPAGEPA 56

Db 1 MGKYMRAKVVVSGEVAAVMEAAAPLGVTRARSLAL-----QKR----- 43

QY 57 AAVSAGDGGSCYIHLRSRMLFMAPPQPSVDSVPTPVEAADGAAGQQAALAGLSRC 116

Db 44 -----QGGG-----YLELRSRRLEKLP PPPP-----PPRRRATAAATAATAA----- 83

QY 117 SSTASSVNLGLGGQSGSHTCRSYDAEA-----GGDHVL-----VDVSAASNSGS 161

Db 84 -----ESAEAEVSFGGENVLEAMEERIAFPVSLNCVKNTWV 120

QY 162 GP-----DRER-----RETPSS-----RAHGELSDLES DLAGH 190

Db 121 APSPKYPRVGATCLOGTKGVPERDEQVLRHFNTRETTPCSLIRDPDTISTPGSTTRS 180

QY 191 KTGPSLPAATPAEELIVPPAHEIQEFFFFAAAEAAQAKRFASKYNFDFVRGVPLDAGGRFEW 250

Db 181 HSSHCKVQTPVRHNIIPASAELEAFAAEQRQAFIDKYNFDPVNDCLPLP--GRFEW 238

RESULT 2

Q7XDH8 PRELIMINARY; PRT; 242 AA.

AC Q7XDH8;

DT 01-OCT-2003 (Tremblrel. 25, Created)

DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE Putative cyclin-dependent kinase inhibitor.

GN OSJNEB0094K03.16.

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzaceae; Oryza.

OX NCBI_TaxID=39947;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Nipponbare;

RA The Rice Chromosome 10 Sequencing Consortium;

RT "In-depth view of structure, activity, and evolution of rice

RT chromosome 10."

RL Science 300:1566-1569(2003).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Nipponbare;

RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;

RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB017104; AAP54233.1; --

KW Cyclin; Kinase.

SQ SEQUENCE 242 AA; 26750 MW; F215A01234735928 CRC64;

Query Match 17.3%; Score 225; DB 10; Length 242;

Best Local Similarity 28.0%; Pred. No. 5.9e-08;

Matches 84; Conservative 24; Mismatches 80; Indels 112; Gaps 12;

QY 1 MGKYMRAKVVVSGEVAAVMEAAAPLGVTRARSLAL-----QKR----- 43

Db 1 MGKYMRAKVVVSGEVAAVMEAAAPLGVTRARSLAL-----QKR----- 43

QY 57 AAVSAGDGGSCYIHLRSRMLFMAPPQPSVDSVPTPVEAADGAAGQQAALAGLSRC 116

Db 44 -----QGGG-----YLELRSRRLEKLP PPPP-----PPRRRATAAATAATAA----- 83

QY 117 SSTASSVNLGLGGQSGSHTCRSYDAEA-----GGDHVL-----VDVSAASNSGS 161

Db 84 -----ESAEAEVSFGGENVLEAMEERIAFPVSLNCVKNTWV 120

QY 162 GP-----DRER-----RETPSS-----RAHGELSDLES DLAGH 190

Db 121 APSPKYPRVGATCLOGTKGVPERDEQVLRHFNTRETTPCSLIRDPDTISTPGSTTRS 180

QY 191 KTGPSLPAATPAEELIVPPAHEIQEFFFFAAAEAAQAKRFASKYNFDFVRGVPLDAGGRFEW 250

Db 181 HSSHCKVQTPVRHNIIPASAELEAFAAEQRQAFIDKYNFDPVNDCLPLP--GRFEW 238

RESULT 3

Q9FKB5 PRELIMINARY; PRT; 222 AA.

AC Q9FKB5;

DT 01-MAR-2001 (Tremblrel. 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE Genomic DNA, chromosome 5, TAC clone:K24G6 (Cyclin-dependent kinase

DE inhibitor 3).

GN KRP3.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Columbia;

RX MEDLINE=98403884; PubMed=9734815;

RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,

RA Tabata S.;

RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.

RT Sequence features of the regions of 1,367,185 bp covered by 19

RT physically assigned P1 and TAC clones.";

RL DNA Res. 5:203-216(1998).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=21342510; PubMed=11449057;

RA de Veylder L., Beeckman T., Beemster G.T.S., Krols L., Terras F.,

RA Landrieu I., Van Der Schueren E., Maes S., Naudts M., Inze D.;

RT "Functional analysis of Cyclin-dependent kinase inhibitors of

RT Arabidopsis.";

RL Plant Cell 13:1653-1668(2001).

DR EMBL; AB012242; BAB09435.1; --

DR EMBL; AJ301554; CAC41617.1; --

DR GO; GO:0005634; C:nucleus; IEA.

DR GO; GO:0004861; F:cyclin-dependent protein kinase inhibitor a. . .; IEA.

DR GO; GO:0016301; F:kinase activity; IEA.

DR GO; GO:0007050; P:cell cycle arrest; IEA.

DR InterPro; IPR003175; CDI.

DR Pfam; PF02234; CDI; 1.

KW Kinase.

SQ SEQUENCE 222 AA; 24925 MW; 7A3EB2C9A29688A7 CRC64;

Query Match 16.3%; Score 212.5; DB 10; Length 222;

Best Local Similarity 29.6%; Pred. No. 3.9e-07;

Matches 79; Conservative 33; Mismatches 94; Indels 61; Gaps 12;

QY 1 MGKYMRAKVVVSGEVAAVMEAAAPLGVTRARSLAL-----QKR----- 43

Db 1 MGKYMRAKVVVSGEVAAVMEAAAPLGVTRARSLAL-----QKR----- 43

QY 57 AAVSAGDGGSCYIHLRSRML-----FMAPPQPSVDSVPTPVEAADGAAGQQAALAA 111

Db 46 AADSALPNDSSCYLQLSRRLKPKPSSLPKQP-PRVH-----RSGIKESG 90

QY 112 GLSRCSTASSVNLGLGGQSGSHTCRSYDAEAAGGDHVLVDVSAASNS-GSGPDRERRET 170

Db 91 SRSRVDSVNS-----VPVAQSSNEDECFDNFVSQVSCGENSLGFESRSTRES 139

QY 171 TPSSRAHAGELSDLES DLAGHKTGPSLPAATP-----AAELIVPPAHEIQEFFFFAAAEAAQAK 226

Db 140 TPCN-----FVEDMEIMVTPGSSSTRSMCRATKEYTREQDNVIPTTSEMEEFFAYAEQQOOR 195

QY 227 RFASKYNFDFVRGVPLDAGGRFEWAPV 253

Db 196 LFMEKYNFDIVNDIPL--SGRYEWVQV 220

RESULT 4

Q8GT28

```
ID Q8GT28 PRELIMINARY; PRT; 185 AA.
AC Q8GT28;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE P27KIP1-related-protein 2.
GN KRP2.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Wval06; TISSUE=Locular;
RA Bisbis B., Joubes J., Hernould M., Inze D., Raymond P., Chevalier C.;
RT "Cyclin-dependent kinase inhibitors are highly expressed in
RT endoreduplicating tissues during the early development of tomato
RT fruits.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ441250; CAD29649.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004861; F:cyclin-dependent protein kinase inhibitor a. . .; IEA.
DR GO; GO:0007050; P:cell cycle arrest; IEA.
DR InterPro; IPR003175; CDI.
DR Pfam; PF02234; CDI; 1.
SQ SEQUENCE 185 AA; 21189 MW; 48DCC89A5336C676 CRC64;

Query Match 15.9%; Score 207; DB 10; Length 185;
Best Local Similarity 27.7%; Pred. No. 7.5e-07;
Matches 70; Conservative 28; Mismatches 85; Indels 70; Gaps 8;

QY 1 MGKYMRCRGAAGAEVAAVEVTQVVGVRTRSRSAATGGVAKVAPRRKRKRAPAGEPAAAVS 60
Db 1 MGKYLKRS-----GHVMESSLGVTRSTILA-----LQLQSSSSSTPPPSLP 44

QY 61 AGDGGGSCYIHLRSMLFMAPPQPPQPSVDSVPTPVEAADGAAGQQAALAAGLSRCSSTA 120
Db 45 SASD--SCYLQLSRRLH-KPPTPIPCNSHPH-----SASV 78

QY 121 SSVNLGLGGQSGSHTCRSYDAAEAGGDHVLVDVSAASNSGSGPDRRERTTPSSRAHGE 180
Db 79 DEISFPDNNLHFQHTRS--TRESTPCSLVREVMVNPFGSATRQTELTNTTQRRNFIL 136

QY 181 SDLESDLAGHKTGPSLPAATPAELIVPPAHEIQEFFFFAAEAAQAKRFASKYNFDFVRGV 240
Db 137 RN-----IPSAHEIEDFFTEAEQQQQLFMDKYNFDDVNDV 172

QY 241 PLDAGGRFEWAPV 253
Db 173 PL--SGRYEWIRV 183

RESULT 5
Q94CM0 PRELIMINARY; PRT; 286 AA.
AC Q94CM0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cyclin-dependent kinase inhibitor 4.
GN KRP4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21342510; PubMed=11449057;
RA de Veylder L., Beeckman T., Beemster G.T.S., Krols L., Terras F.,
RA Landrieu I., Van Der Schueren E., Maes S., Naudts M., Inze D.;
RT "Functional analysis of Cyclin-dependent kinase inhibitors of
```

```
RT Arabidopsis.";
RL Plant Cell 13:1653-1668(2001).
DR EMBL; AJ301555; CAC41618.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004861; F:cyclin-dependent protein kinase inhibitor a. . .; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0007050; P:cell cycle arrest; IEA.
DR InterPro; IPR003175; CDI.
DR Pfam; PF02234; CDI; 1.
KW Kinase.
SQ SEQUENCE 286 AA; 31719 MW; 1CEB472583302AAD CRC64;

Query Match 13.8%; Score 179.5; DB 10; Length 286;
Best Local Similarity 24.1%; Pred. No. 9.3e-05;
Matches 81; Conservative 33; Mismatches 87; Indels 135; Gaps 15;

QY 1 MGKYMRCR-----GAAG-----AEVAAVEV-----TQVVGVRTRSRSAATGGV 40
Db 1 MGKYIRKSKIDGAGAGAGGGGGGGGESSIALMDVVSPSSSSSLGVLTRAKSLALQQQQ 60

QY 41 AKVAPRRKRKRAPAGEPAAAVSAG-----GDGGSCYIHLRSMLFMAP----- 81
Db 61 QRCLLQKPSSPSSLPPTSASPNPPSKQMKKKQQQMNDGGS-YLQLRSRRLQKKPPIVVI 119

QY 82 -----PQPPSVDSVPTPVEAADGAAGQQAALAAAGLSRCSSTASSV 123
Db 120 RSTKRRKQQRNETCGRNPNSNLDSE-----RGDG-----SRSDSVSESV 161

QY 124 NLGLGGQSGSHTCRS-----YDAEAGGDHVLVDVSAASNSGSGPDRRERTTPSSRA 176
Db 162 VFGKDKDLISEINKDPTFGQNFLEE---EHT-----QRTTRESTPCSII 204

QY 177 H-----GELSDLESCLA-GHKTGPSLPAATPAELIVPPAHEIQEFFF 217
Db 205 RRPEIMTTPGSSSTKINICVSESNQREDLSRSHRRRPTTP-----EMDEFF 250

QY 218 AAAEAAQAKRFASKYNFDFVRGVPLDAGGRFEWAPV 253
Db 251 SGAEEBQQKQFIEKYNFDPVNEQPLP--GRFEWTKV 284

RESULT 6
O48597 PRELIMINARY; PRT; 196 AA.
AC O48597;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cyclin-dependent kinase inhibitor protein.
GN CDK11.
OS Chenopodium rubrum (Red goosefoot) (Pigweed).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Amaranthaceae; Chenopodium.
OX NCBI_TaxID=3560;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93329903; PubMed=10409053;
RA Fountain M.D., Renz A., Beck E.;
RT "Isolation of a cDNA encoding a G1-cyclin-dependent kinase inhibitor
RT from suspension-cultured photoautotrophic chenopodium rubrum cells.";
RL Plant Physiol. 120:339-339(1999).
DR EMBL; AJ002173; CAA05215.1; -.
DR PIR; T09968; T09968.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004861; F:cyclin-dependent protein kinase inhibitor a. . .; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0007050; P:cell cycle arrest; IEA.
DR InterPro; IPR003175; CDI.
DR Pfam; PF02234; CDI; 1.
KW Kinase.
SQ SEQUENCE 196 AA; 21685 MW; 150349F7ED82A6B2 CRC64;
```


QY 124 NLGLGGQSGSHTCRS-----YDAAGAGDHDVLDVSAASNSGSGPDRETTTPSSRA 176
Db 162 VFGKDKDLISEINKDPTFGQNFDFLEE---EHT-----QRTTRESTPCSLI 204
QY 177 H-----GELSDLESDLA-GHKTGPSLPAATPAALIVPPAHEIQEFF 217
Db 205 RRPEIMTTPGSSSTKLNICVSESQREDLSRSHRRPTTP-----EMGEFF 250
QY 218 AAAAQAQAKRFASKYNFDFVRGVPLDAGGRFEWAPV 253
Db 251 SGAEEQKQKQFIEKYNFDPVNEQPLP--GRFEWTKV 284
RESULT 9
Q9FS28
ID Q9FS28 PRELIMINARY; PRT; 192 AA.
AC Q9FS28;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cyclin dependent kinase inhibitor.
GN CKI.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Axillary bud;
RA Shimizu-Sato S., Mori H.;
RT "A cDNA from Pisum sativum encoding the cyclin-dependent kinase
inhibitor (CKI) homologue";
RL Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB029483; BAB20860.1; -;
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004861; F:cyclin-dependent protein kinase inhibitor a. . .; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0007050; P:cell cycle arrest; IEA.
DR InterPro; IPR003175; CDI.
DR Pfam; PF02234; CDI; 1.
KW Kinase.
SQ SEQUENCE 192 AA; 21921 MW; E76D734000885EB2 CRC64;
Query Match 13.0%; Score 170; DB 10; Length 192;
Best Local Similarity 27.8%; Pred. No. 0.00027;
Matches 65; Conservative 28; Mismatches 81; Indels 60; Gaps 11;
QY 25 VGVTRRSRA---AATGGVAKVAPRRKRAPAGEPAAAVSAGDGGGSCYIHLRSRMLFMAP 81
Db 6 VGVTRARAALAMEATSSPRTTKRK-----INRTENRKFS- 43
QY 82 PQQPSVDSV-PTPVEAADGAAGQQAALAGLSRCSSTASSVNLGLGGQSGSHTCRSYD 140
Db 44 --VKPKIATVRPETVTEKHSSGSTSDEEFPAS--CCSSNGSVEL-----DEERIKSLD 92
QY 141 AAAGAGDHDVLDVSA---SNSGSGPDRETTTPSSRAHAGLSLDSLDLGHKTGPSL 196
Db 93 -----LEVESAQGETSTCNCDEEIEREMSRSSSEFRGNSHELES-----METNSRR 138
QY 197 PAATPAALIVPPAHEIQEFFAAAEAAQAKRFASKYNFDFVRGVPLDAGGRFEW 250
Db 139 PISSPKK---TPTEYELEEEFAAAEKDIQKFKQKYNIDILKDVPLE--GRYEW 187
RESULT 10
Q8GT29
ID Q8GT29 PRELIMINARY; PRT; 210 AA.
AC Q8GT29;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE P27KIP1-related-protein 1.
GN KRPI.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Wva106; TISSUE=Locular;
RA Bisbis B., Joubes J., Hernould M., Inze D., Raymond P., Chevalier C.;
RT "Cyclin-dependent kinase inhibitors are highly expressed in
endoreduplicating tissues during the early development of tomato
fruits.";
RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ441249; CAD29648.1; -;
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004861; F:cyclin-dependent protein kinase inhibitor a. . .; IEA.
DR GO; GO:0007050; P:cell cycle arrest; IEA.
DR InterPro; IPR003175; CDI.
DR Pfam; PF02234; CDI; 1.
SQ SEQUENCE 210 AA; 23845 MW; 38FB15C35044FED6 CRC64;
Query Match 13.0%; Score 169.5; DB 10; Length 210;
Best Local Similarity 26.4%; Pred. No. 0.00032;
Matches 70; Conservative 23; Mismatches 103; Indels 69; Gaps 8;
QY 1 MGKYMRCRGAGAEVAAVEVTQVVGVRTSRSAATGGVAKVAPRRKRAPAGEPAAAVS 60
Db 1 MGKYIRKTR-----KTEDVSPGLVTRAKALALN----- 29
QY 61 AGDGGGSCYIHLRSRMLFMAPQPOPSVDSVPTPVEAADGAAGQQAALAGLSRCSSTA 120
Db 30 -GGDGGG-YLELRSRLV-----KPTVLEGRQKNGVKNPNLVPNPQ 73
QY 121 SSVNLGLGGQSGSHTCRSYDAEAGGDHVLVDVSAASN--SGSGPDRETTTPSSRAHG 178
Db 74 QIPNVCVNSEEGKGVKEMENQKEKSCLGPEDSFGENLLEFEGRKRTTRESTPCS---- 129
QY 179 ELSDLESDLGHKTGPSLPAATPAAL-----IVPPAHEIQEFFAAAEAAQAKRF 228
Db 130 ----LIRDSDNIQTPGSSTRRTNANEANGRVPSNIQPTIPTDLEMEEFFTRAKEQQRKF 185
QY 229 ASKYNFDFVRGVPLDAGGRFEWAPV 253
Db 186 IEKYNFDPVNEKPLP--GRYEWVKV 208
RESULT 11
Q8GUA2
ID Q8GUA2 PRELIMINARY; PRT; 207 AA.
AC Q8GUA2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cyclin-dependent kinase inhibitor.
GN KIS2.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RA Jasinski S., Saraiva Leite C., Domenichini S., Stevens R., Raynaud C.,
RA Perennes C., Bergounioux C., Glab N.;
RT "NtKIS2, a novel tobacco cyclin-dependent kinase inhibitor
differentially expressed during the cell cycle and plant
development.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AJ517189; CAD56868.1; -;
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004861; F:cyclin-dependent protein kinase inhibitor a. . .; IEA.

Db 47 -----CRSH-----SGVDTTPAGSLV-----SPSS 66
QY 122 SVNGLGGQSGHTCRSYDAAEAGGDHVLVDVSAASNGS-GPDRERETTPSSRAHGE 180
Db 67 SVNLN-DASNLHDLASCYLRNGSSENSVIAESAESKEAKLSERQR--TPEK----- 116
QY 181 SDLESDLGHKGTGSLPAATPAELIVPPAHEIQEFFFFAAAEAAQAKRFASKYNFDFVRGV 240
Db 117 -----MPSEKEIEEFAARQKAILKRFRKKYNFDEPEKEE 150
QY 241 PLDAGGRFEWAPVVS 255
Db 151 PLE--GRYEWVRIGS 163

RESULT 14

Q9LJL5
ID Q9LJL5 PRELIMINARY; PRT; 196 AA.
AC Q9LJL5;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Gb|AAC23758.1 (Cyclin-dependent kinase inhibitor 5).
GN KRP6 OR ACK1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=20363099; PubMed=10907853;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety P1,
RT TAC and BAC clones."
RL DNA Res. 7:217-221(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21342510; PubMed=11449057;
RA de Veylder L., Beeckman T., Beemster G.T.S., Krols L., Terras F.,
RA Landrieu I., Van Der Schueren E., Maes S., Naudts M., Inze D.;
RT "Functional analysis of Cyclin-dependent kinase inhibitors of
RT Arabidopsis."
RL Plant Cell 13:1653-1668(2001).
RN [4]
RP SEQUENCE FROM N.A.
RA Park S.C., Cho J.W.;
RT "A novel cyclin dependent kinase inhibitor (ack1).";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP000419; BAB02955.1; --
DR EMBL; AJ301557; CAC41620.1; --
DR EMBL; AF106705; AAF77612.2; --
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004861; F:cyclin-dependent protein kinase inhibitor a. . .; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0007050; P:cell cycle arrest; IEA.
DR InterPro; IPR003175; CDI.
DR Pfam; PF02234; CDI; 1.
KW Cyclin; Kinase.
SQ SEQUENCE 196 AA; 21454 MW; AC543304CE6B92C4 CRC64;

Query Match 10.3%; Score 134; DB 10; Length 196;
Best Local Similarity 27.2%; Pred. No. 0.08;
Matches 52; Conservative 26; Mismatches 57; Indels 56; Gaps 9;

QY 87 SVDSVPTPVEAADGAAGQQAALAGLSRCSSTASSVNLGLGQRG-----SHTCRSYD 140
Db 32 SPDSHDVIVFAVSSSVASSAALAS--DECSVT-----IGEESDQSSSISSGCTTSE 82
QY 141 AAEAGGDHVLVDVSAASNGSGPDRE-----RRETPSSRAHGE-LS 181
Db 83 SKEI-----AKNSSFGVDLEDHQIETETETSTFITSNFKETSPVSEGLGETT 132
QY 182 DLESDLGHKGTGSLPAATPAELIVPPAHEIQEFFFFAAAEAA--QAKRFASKYNFDFVRG 239
Db 133 EMESSATKRKQPGVRK-----TPTAAEIEDLFSELESQDDKKKQFIEKYNFEDIVND 184
QY 240 VPLDAGGRFEW 250
Db 185 EPLE--GRYKW 193

RESULT 15

Q8S9U7
ID Q8S9U7 PRELIMINARY; PRT; 204 AA.
AC Q8S9U7;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Similar to cyclin-dependent kinase inhibitor (Putative cyclin-
DE dependent kinase inhibitor).
GN B1064G04.1 OR B1144D11.17.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC
RT clone:B1064G04.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC
RT clone:B1144D11.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003924; BAB85380.1; --
DR EMBL; AP003443; BAC00670.1; --
DR Gramene; Q8S9U7; --
DR GO; GO:0016301; F:kinase activity; IEA.
KW Cyclin; Kinase.
SQ SEQUENCE 204 AA; 21233 MW; 900E2DDAB299150 CRC64;

Query Match 9.9%; Score 129.5; DB 10; Length 204;
Best Local Similarity 27.5%; Pred. No. 0.17;
Matches 61; Conservative 18; Mismatches 50; Indels 93; Gaps 12;

QY 1 MGKYMRCRCGAAGA--EVAAVEVTQVVGVTRTSRRAAATGGVAKVAPRRKRAPAGEPAAA 58
Db 1 MGKYMRAKVAAAVMELAAV-APAPLGVTRTRSLAL-----QKR----- 40
QY 59 VSAGDGGSCYIHLRSRMLFMAPQPPQPSVDSVP---TPVEAADGAAGQQAALAGLSR 115
Db 41 --QGGE---YLELRSRRLKLPPLPPPPP---PRRRAPVTTDPAAA----- 78
QY 116 CSSTASSVNLGLGGQSGHTCRSYDAAEAGGDHVLVDVSAASNGSGPDRETTTPSSR 175
Db 79 -----ESAEAG-----SFGGGER-----PRAG 97
QY 176 AHGELSDLESDLGHKGTGPSL--PAATPAELIV---PPAHE 212
Db 98 GHGKVQSLADQAPGITGGVLDPPAAKHATGTVADRRRPLHK 139

Search completed: October 1, 2004, 16:53:48
Job time : 119 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 16:38:10 ; Search time 121 Seconds
(without alignments)
597.786 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 1304
Sequence: 1 MGKYMRCRGAGAEVAAVE.....VRGVPLDAGGRFEWAPVVSII 256

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 1304 | 100.0 | 256 | 6 | ABB98757 Maize CKI |
| 2 | 669 | 51.3 | 262 | 5 | ABG65670 OSICK 2 p |
| 3 | 665 | 51.0 | 227 | 6 | ABB98759 Maize CKI |
| 4 | 612 | 46.9 | 125 | 3 | AAB26250 Corn cycl |
| 5 | 411 | 31.5 | 126 | 3 | AAB26249 Corn cycl |
| 6 | 399.5 | 30.6 | 116 | 5 | ABG65674 Corn ICK |
| 7 | 279 | 21.4 | 226 | 5 | ABG65692 Rice OsIC |
| 8 | 235 | 18.0 | 194 | 5 | ABG65691 Rice OsIC |
| 9 | 232.5 | 17.8 | 53 | 3 | AAB26245 Corn cycl |
| 10 | 232.5 | 17.8 | 53 | 3 | AAP01940 Corn cycl |
| 11 | 219 | 16.8 | 190 | 3 | AAB27258 Wheat cyc |
| 12 | 217 | 16.6 | 42 | 6 | ABB98860 Maize CKI |
| 13 | 215.5 | 16.5 | 222 | 5 | AAE25111 Arabidops |
| 14 | 204 | 15.6 | 190 | 3 | AAB26252 Corn cycl |
| 15 | 202.5 | 15.5 | 46 | 6 | ABB98862 Maize CKI |
| 16 | 200 | 15.3 | 190 | 6 | ABB98758 Maize CKI |
| 17 | 190.5 | 14.6 | 216 | 2 | AAW98180 Arabidops |
| 18 | 178 | 13.7 | 196 | 3 | AAB27259 Chenopodi |
| 19 | 178 | 13.7 | 196 | 3 | AAV44340 Chenopodi |
| 20 | 178 | 13.7 | 196 | 5 | AAE25108 Chenopodi |
| 21 | 178 | 13.7 | 289 | 3 | AAB27262 Arabidops |
| 22 | 178 | 13.7 | 289 | 3 | AAG40485 Arabidops |
| 23 | 178 | 13.7 | 289 | 5 | AAE25109 Arabidops |
| 24 | 178 | 13.7 | 289 | 5 | AAU72581 Arabidops |
| 25 | 160 | 12.3 | 205 | 3 | AAP01952 Cyclin de |

| | | | | | | |
|----|-------|------|-----|---|-----------|-----------|
| 26 | 160 | 12.3 | 205 | 3 | AAB27254 | Soybean c |
| 27 | 156 | 12.0 | 256 | 3 | AAG40486 | Arabidops |
| 28 | 149 | 11.4 | 327 | 3 | AAB27261 | Arabidops |
| 29 | 147.5 | 11.3 | 136 | 3 | AAAY44339 | Arabidops |
| 30 | 145.5 | 11.2 | 37 | 3 | AAP01950 | Rice cycl |
| 31 | 145.5 | 11.2 | 37 | 3 | AAB27252 | Rice cycl |
| 32 | 145 | 11.1 | 137 | 5 | AAE25107 | Arabidops |
| 33 | 144.5 | 11.1 | 108 | 5 | ABG65676 | Broom cor |
| 34 | 142.5 | 10.9 | 54 | 3 | AAB26253 | Corn cycl |
| 35 | 137 | 10.5 | 201 | 3 | AAG40487 | Arabidops |
| 36 | 134 | 10.3 | 90 | 5 | ABG65673 | Rice Os I |
| 37 | 134 | 10.3 | 99 | 3 | AAB27251 | Rice cycl |
| 38 | 134 | 10.3 | 99 | 3 | AAP01949 | Cyclin de |
| 39 | 134 | 10.3 | 196 | 5 | AAE25104 | Arabidops |
| 40 | 134 | 10.3 | 208 | 4 | AAB47001 | Plant D-1 |
| 41 | 134 | 10.3 | 212 | 3 | AAAY44337 | Arabidops |
| 42 | 134 | 10.3 | 212 | 5 | AAE25103 | Arabidops |
| 43 | 129.5 | 9.9 | 93 | 5 | ABG65675 | Corn ICK |
| 44 | 129 | 9.9 | 88 | 3 | AAB27257 | Soybean c |
| 45 | 126 | 9.7 | 60 | 3 | AAP01953 | Cyclin de |

ALIGNMENTS

RESULT 1
ABB98757
ID ABB98757 standard; protein; 256 AA.
XX AC ABB98757;
XX AC ABB98757;
DT 21-FEB-2003 (first entry)
XX Maize CKI_B.
DE Maize CKI_B.
XX Maize; cyclin-dependent kinase inhibitor; CKI; CKI B; plant; corn;
KW crop yield; root size; plant growth; tassel size; ear size;
KW male sterility; endoreduplication.
XX Zea mays.
OS Zea mays.
XX Zea mays.
FH Key Location/Qualifiers
FT Domain 1..7
FT /note= "Conserved domain"
XX WO200281623-A2.
XX 17-OCT-2002.
XX 06-NOV-2001; 2001WO-US044038.
XX 07-NOV-2000; 2000US-0246349P.
XX (PION-) PIONEER HI-BRED INT INC.
XX (ARIZ-) ARIZONA BOARD OF REGENTS.
XX Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;
XX WPI; 2003-058511/05.
XX N-PSDB; ABV74603.
XX Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
XX involved in cell cycle regulation, and useful for altering cell cycle
XX protein content, cell cycle progression, cell number and composition of
XX plants.
XX Claim 12; Page 65-66; 69pp; English.
XX The present sequence is maize cyclin-dependent kinase inhibitor (CKI),
XX CKI_B. The coding sequence for this protein (i) is useful for modulating
XX the activity of cyclin-dependent kinase (CDK) in a plant such as a corn,
XX soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice,
XX barley, oil-seed Brassica and millet. Modulating the activity of CDK,

XX 07-NOV-2000; 2000US-0246349P.
PR (PION-) PIONEER HI-BRED INT INC.
XX (ARIZ-) ARIZONA BOARD OF REGENTS.
PA Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;
XX WPI; 2003-058511/05.
XX N-PSDB; ABV74605.
XX Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
PT involved in cell cycle regulation, and useful for altering cell cycle
PT protein content, cell cycle progression, cell number and composition of
PT plants.
XX Claim 12; Page 69; 69pp; English.
PS The present sequence is maize cyclin-dependent kinase inhibitor (CKI),
XX CKI_D. The coding sequence for this protein (I) is useful for modulating
CC the activity of cyclin-dependent kinase (CDK) in a plant such as a corn,
CC soybean, sunflower, sorghum, canola, alfalfa, cotton, rice,
CC barley, oil-seed Brassica and millet. Modulating the activity of CDK,
CC preferably modulating downward is useful for providing differential
CC growth in a plant, especially a positive growth advantage and modulating
CC CDK activity upward is useful for increasing crop yield, root size, plant
CC growth, tassel size and/or ear size. Modulating CDK activity is also
CC useful for conferring male sterility and for improving transformation
CC frequencies by increasing the number of cells in cell division. CDK
CC activity can also be modulated for modulating endoreduplication in the
CC endosperm of corn, sorghum, wheat, rice, barley, and millet, where the
CC promoter used is an endosperm-preferred promoter. The cell numbers are
CC modulated in one or more tissues of a plant, comprising root, seed,
CC tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves,
CC stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm,
CC pericarp, meristematic or leaf cells. (I) is also useful for identifying
CC maize CKI interacting proteins, by aducting the nucleic acid sequence to
CC a second nucleic acid sequence encoding a DNA-binding domain
XX
SQ Sequence 227 AA;
Query Match 51.0%; Score 665; DB 6; Length 227;
Best Local Similarity 72.6%; Pred. No. 9.3e-52;
Matches 172; Conservative 6; Mismatches 31; Indels 28; Gaps 16;
QY 1 MGKYMRCRGAAGAEVAAVEVTQVVGVTFRSRAAAT-GGVAKVA-PRRKRA--PA---- 52
Db 1 MGKYMRCRGAAGAEVAAVEVSQVVGVTFRSRAAATGGGVAKVAPRRKKALLPAANVT 60
QY 53 --GEPAAVSAAGDGGG-CYIHLRSMLFMAPPQPPQPSVDSVPTPVEAADGAAGQOG-AA 108
Db 61 TSGEPGAVGAGDGGGCCYIHLRSMLFMAPPQPPQPS--AALTPEAA-GAAXQGGVVA 117
QY 109 LAAGLSRCSSTASSVNLGLGGQSGHTCSYDA-AEAGGDHVLVDVSAASNSGSGPDRER 167
Db 118 LAAGLSRCSSTASSVDV-----GGHACRSDAAPAEVDGDHV-PDVVTASNSGSGVPDRER 170
QY 168 RETTP-SSRAH-GELSDLESLAG-HKTG-PSLPA-ATPAELIVPPAHEIQEFFAA 219
Db 171 RETTPSSRAHGGELSDLESLVGRQKTGCSSTPATTTSAAELIVPPAHEIQEFFAA 227
RESULT 4
AAB26250
ID AAB26250 standard; protein; 125 AA.
XX
AC AAB26250;
XX
DT 17-JAN-2001 (first entry)
XX
DE Corn cyclin-dependent kinase inhibitor #3.
XX
KW Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;

KW cell growth; herbicide.
XX
OS Zea mays.
XX
PN WO200060087-A2.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US009106.
XX
PR 07-APR-1999; 99US-0128192P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Klein TM, Weng Z, Cahoon RE;
XX
DR WPI; 2000-679375/66.
DR N-PSDB; AAA95281.
XX
PT Cyclin dependent kinase inhibitor sequences, useful for identifying
PT herbicides and plant growth inhibitors.
XX
PS Claim 10; Fig 1; 58pp; English.
XX
CC The present sequence is the corn cyclin-dependent kinase inhibitor
CC (CDKI). Its coding sequence was isolated by searching a contig comprising
CC cDNA from corn roots and ear leaf sheath for sequences similar to those
CC encoding the CDKI from Chenopodium rubrum, Caenorhabditis elegans and
CC Arabidopsis thaliana. CDKI is involved in the cell cycle, and may promote
CC or inhibit cell division and growth. The protein and its coding sequence
CC are useful in the production of transgenic plants which produce increased
CC or decreased amounts of the CDKI protein, in the identification of
CC herbicides, in genetic and physical mapping and in the isolation of the
CC CDKI gene in other organisms
XX
SQ Sequence 125 AA;
Query Match 46.9%; Score 612; DB 3; Length 125;
Best Local Similarity 96.0%; Pred. No. 2.6e-47;
Matches 120; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 132 GSHTCRSYDAAEAGGDHVLVDVSAASNSGSGPDRERETTPSSRAHGEISDLESLAGHK 191
Db 1 GSHTCRSDDAEAGGDHVLVDVSAASNSGSGPDRERETTPSSRAHGEISDLESLAGHK 60
QY 192 TGPSLPAAATPAAELIVPPAHEIQEFFAAAEAAQAKRFASKYNFDFVRGVPLDAGGRFEWA 251
Db 61 TGPSLPAAATPAAELIVPPAHEIQEFFAAAEAAQAKRFASKYNFDFVRGVPLDAGGRFEWA 120
QY 252 PVVSI 256
Db 121 PVVSI 125
RESULT 5
AAB26249
ID AAB26249 standard; protein; 126 AA.
XX
AC AAB26249;
XX
DT 17-JAN-2001 (first entry)
XX
DE Corn cyclin-dependent kinase inhibitor #2.
XX
KW Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
KW cell growth; herbicide.
XX
OS Zea mays.
XX
PN WO200060087-A2.
XX
PD 12-OCT-2000.
XX

PF 06-APR-2000; 2000WO-US009106.
XX
PR 07-APR-1999; 99US-0128192P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Klein TM, Weng Z, Cahoon RE;
XX
DR WPI; 2000-679375/66.
DR N-PSDB; AAA95280.
XX
PT Cyclin dependent kinase inhibitor sequences, useful for identifying
PT herbicides and plant growth inhibitors.
XX
PS Claim 10; Fig 1; 58pp; English.
XX
CC The present sequence is the corn cyclin-dependent kinase inhibitor
CC (CDKI). Its coding sequence was isolated by searching a corn silk cDNA
CC library for sequences similar to those encoding the CDKI from Chenopodium
CC rubrum, Caenorhabditis elegans and Arabidopsis thaliana. CDKI is involved
CC in the cell cycle, and may promote or inhibit cell division and growth.
CC The protein and its coding sequence are useful in the production of
CC transgenic plants which produce increased or decreased amounts of the
CC CDKI protein, in the identification of herbicides, in genetic and
CC physical mapping and in the isolation of the CDKI gene in other organisms
XX
SQ Sequence 126 AA;

Query Match 31.5%; Score 411; DB 3; Length 126;
Best Local Similarity 78.7%; Pred. No. 3.6e-29;
Matches 100; Conservative 1; Mismatches 18; Indels 8; Gaps 8;

QY 136 CRSYDA-AEAGGDHVLVDVSAASNSGSGPDRERETTP-SSRAH-GELSDLESDLAG-HK 191
Db 2 CRSDAAPAEVDGDHV-PDVVTASNSGSGVPDRERETTPSSSRAHGGELSDLESDLVGRQK 60

QY 192 TG-PSLPAATPA-AELIVPPAHEIQEFAAAEAAQAKRFASKYNFDFVRGVPLDAGGRFE 249
Db 61 TGCSSSPATTTSVAELIVPPAQEIQEFFAAEAAHAKRFASKYNFDFVRGVPLDA-GRFE 119

QY 250 WAPVVS I 256
Db 120 WTPGVSI 126

RESULT 6
ABG65674
ID ABG65674 standard; protein; 116 AA.
XX
AC ABG65674;
XX
DT 07-AUG-2003 (revised)
DT 27-AUG-2002 (first entry)
XX
DE Corn ICK 1 protein.
XX
KW Plant; inhibitor of cyclin dependent kinase; ICK.
XX
OS Zea mays.
XX
PN WO200228893-A2.
XX
PD 11-APR-2002.
XX
PF 29-JUN-2001; 2001WO-IB001492.
XX
PR 14-JUL-2000; 2000US-0218471P.
PR 13-OCT-2000; 2000US-0241219P.
XX
PA (CROP-) CROPDESIGN NV.
XX
PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
PI Hatzfeld Y;

XX WPI; 2002-471311/50.
DR N-PSDB; ABK93954.
XX
PT Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
PT to screen substrates, drugs or compounds which modulate ICK activity and
PT treat disorders characterized by an insufficient or excessive production
PT of ICK inhibitors.
XX
PS Claim 48; Fig 4; 141pp; English.
XX
CC This invention relates to the DNA and protein sequences of novel isolated
CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
CC the invention may be used for treating disorders characterised by
CC insufficient or excessive production of an ICK inhibitor. The protein of
CC the invention may also be used to screen for naturally-occurring ICK
CC substrates, drugs or compounds which modulate ICK activity, as well as to
CC treat disorders characterised by insufficient or excessive production of
CC ICK protein, forms which have decreased or aberrant activity compared to
CC ICK wild type protein. The present sequence represents an inhibitor of
CC cyclin-dependent kinase (ICK) protein of the invention. (Updated on 07-
CC AUG-2003 to correct OS field.)
XX
SQ Sequence 116 AA;

Query Match 30.6%; Score 399.5; DB 5; Length 116;
Best Local Similarity 81.9%; Pred. No. 3.6e-28;
Matches 95; Conservative 0; Mismatches 14; Indels 7; Gaps 7;

QY 146 GDHVLVDVSAASNSGSGPDRERETTP-SSRAH-GELSDLESDLAG-HKTG-PSLPA-AT 200
Db 3 GDHV-PDVVXASNSGSGVPDRERETTPSSSRAHGGELSDLESDLVGRQKTGSSSPATT 61

QY 201 PAAELIVPPAHEIQEFAAAEAAQAKRFASKYNFDFVRGVPLDAGGRFEWAPVVS I 256
Db 62 SAAELIVPPAQEIQEFFAAEAAHAKRFASKYNFDFVRGVPLDA-GRFEWTPGVSI 116

RESULT 7
ABG65692
ID ABG65692 standard; protein; 226 AA.
XX
AC ABG65692;
XX
DT 27-AUG-2002 (first entry)
XX
DE Rice OsICK5 protein.
XX
KW Plant; inhibitor of cyclin dependent kinase; ICK.
XX
OS Oryza sativa.
XX
PN WO200228893-A2.
XX
PD 11-APR-2002.
XX
PF 29-JUN-2001; 2001WO-IB001492.
XX
PR 14-JUL-2000; 2000US-0218471P.
PR 13-OCT-2000; 2000US-0241219P.
XX
PA (CROP-) CROPDESIGN NV.
XX
PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
PI Hatzfeld Y;
XX
DR WPI; 2002-471311/50.
DR N-PSDB; ABK93981.
XX
PT Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
PT to screen substrates, drugs or compounds which modulate ICK activity and
PT treat disorders characterized by an insufficient or excessive production
PT of ICK inhibitors.

Cyclin dependent kinase inhibitor sequences, useful for identifying herbicides and plant growth inhibitors.

Claim 10; Page 40; 58pp; English.

The present sequence is the corn cyclin-dependent kinase inhibitor (CDKI). Its coding sequence was isolated by searching a corn silk cDNA library for sequences similar to those encoding the CDKI from *Chenopodium rubrum*, *Caenorhabditis elegans* and *Arabidopsis thaliana*. CDKI is involved in the cell cycle, and may promote or inhibit cell division and growth. The protein and its coding sequence are useful in the production of transgenic plants which produce increased or decreased amounts of the CDKI protein, in the identification of herbicides, in genetic and physical mapping and in the isolation of the CDKI gene in other organisms

Sequence 53 AA;

```

CC markers. The sequence is the corn CDKI clone csiln.pk0050.e6 as described
CC in the method of the invention
XX
SQ Sequence 53 AA;

Query Match 17.8%; Score 232.5; DB 3; Length 53;
Best Local Similarity 88.7%; Pred. No. 1.5e-13;
Matches 47; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 200 TPAAEILVPPAHEIQEFFFFAAAEAAQAKRFASKYNFDFVRGVPLDAGGRFEWAP 252
Db 2 TSVAEILVPPAQETQEFFFFAAAEAAHAKRFASKYNFDFVRGVPLDA-GRFEWTP 53

RESULT 11
AAB27258
ID AAB27258 standard; protein; 190 AA.
XX AC
XX AAB27258;
XX
DT 17-JAN-2001 (first entry)
XX
DE Wheat cyclin-dependent kinase inhibitor #2.
XX
KW Wheat; cyclin-dependent kinase inhibitor; cell cycle; cell division;
KW CDKI; cell growth; herbicide.
XX
OS Triticum aestivum.
XX
PN WO200060087-A2.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US009106.
XX
PR 07-APR-1999; 99US-0128192P.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI Klein TM, Weng Z, Cahoon RE;
XX
XX WPI; 2000-679375/66.
DR N-PSDB; AAA95292.
XX
XX Cyclin dependent kinase inhibitor sequences, useful for identifying
PT herbicides and plant growth inhibitors.
XX
PS Claim 10; Fig 1; 58pp; English.
XX
XX The present sequence is the wheat cyclin-dependent kinase inhibitor
CC (CDKI). Its coding sequence was isolated by searching a wheat root cDNA
CC library for sequences similar to those encoding the CDKI from Chenopodium
CC rubrum, Caenorhabditis elegans and Arabidopsis thaliana. CDKI is involved
CC in the cell cycle, and may promote or inhibit cell division and growth.
CC The protein and its coding sequence are useful in the production of
CC transgenic plants which produce increased or decreased amounts of the
CC CDKI protein, in the identification of herbicides, in genetic and
CC physical mapping and in the isolation of the CDKI gene in other organisms
XX
SQ Sequence 190 AA;

```

Claim 10; Page 40; 58pp; English.

The invention describes a novel isolated polynucleotide comprising a nucleotide sequence encoding one of 17 specific cyclin dependent kinase inhibitor (CDKI) polypeptides, cell cycle regulators involved in control of cell division, growth and death. The nucleotide sequences can be used in a vector to transform a host cell to produce the CDKI polypeptide. They can also be used in methods for selecting and obtaining a nucleic acid sequence that encodes CDKI or affects the level of CDKI expression. The encoded protein can be used in a method for evaluating a compound for its ability to inhibit the activity of a CDKI. The inhibitors can be used as herbicides. They can also be used to inhibit plant growth. The polynucleotide sequences can be used in gene mapping and as genetic

QY 120 ASSVNLGLGQSGHSTCRSYDAAEAGGDHVLVDVSAASNSGSGPDRERRETPSS--RAH 177
Db 76 AERVE-----AEAEDEVSGFNVLESEAMG--RGTRTPCSLIRDS 116
QY 178 GELSDLESLAGHKTGPS-----LPAATPAEELIVPPAHEIQEFAAAAEAAQAKRFASKY 232
Db 117 GTIS-----TPGSTTRPSHNSHRRVQAPARH-IIPCSAEMNEFFSAEQPQQAFIDKY 170
QY 233 NFDFVRGVPLDAGGRFEW 250
Db 171 NFDPVNDCLP--GRYEW 186

RESULT 12
ABB98860
ID ABB98860 standard; protein; 42 AA.
XX
AC ABB98860;
XX
DT 21-FEB-2003 (first entry)
XX
DE Maize CKI_B C-terminal protein fragment.
XX
KW Maize; cyclin-dependent kinase inhibitor; CKI; plant; corn; crop yield;
KW root size; plant growth; tassel size; ear size; male sterility;
KW endoreduplication.
XX
OS Zea mays.
XX
PN WO200281623-A2.
XX
PD 17-OCT-2002.
XX
PF 06-NOV-2001; 2001WO-US044038.
XX
PR 07-NOV-2000; 2000US-0246349P.
XX
XX (PION-) PIONEER HI-BRED INT INC.
PA (ARIZ-) ARIZONA BOARD OF REGENTS.
XX
PI Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;
XX WPI; 2003-058511/05.
DR
XX
PT Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
PT involved in cell cycle regulation, and useful for altering cell cycle
PT protein content, cell cycle progression, cell number and composition of
PT plants.

PS Example 11; Page 46; 69pp; English.
XX
CC The present invention relates to maize cyclin-dependent kinase inhibitor
CC (CKI), CKI_B, CKI_C and CKI_D. The coding sequences (I) for these
CC proteins are useful for modulating the activity of CDK in a plant such as
CC a corn, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton,
CC rice, barley, oil-seed Brassica and millet. Modulating the activity of
CC CDK, preferably modulating downward is useful for providing differential
CC growth in a plant, especially a positive growth advantage and modulating
CC CDK activity upward is useful for increasing crop yield, root size, plant
CC growth, tassel size and/or ear size. Modulating CDK activity is also
CC useful for conferring male sterility and for improving transformation
CC frequencies by increasing the number of cells in cell division. CDK
CC activity can also be modulated for modulating endoreduplication in the
CC endosperm of corn, sorghum, wheat, rice, barley, and millet, where the
CC promoter used is an endosperm-preferred promoter. The cell numbers are
CC modulated in one or more tissues of a plant, comprising root, seed,
CC tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves,
CC stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm,
CC pericarp, meristematic or leaf cells. (I) is also useful for identifying
CC maize CKI interacting proteins, by adducting the nucleic acid sequence to
CC a second nucleic acid sequence encoding a DNA-binding domain. The present
CC sequence is a C-terminal protein fragment of CKI_B. This sequence
CC contains a CDK binding region and/or cyclin binding domains

XX
SQ Sequence 42 AA;
Query Match 16.6%; Score 217; DB 6; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.8e-12;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 215 EFAAAEAAQAKRFASKYNFDFVRGVPLDAGGRFEWAPVSI 256
Db 1 EFAAAEAAQAKRFASKYNFDFVRGVPLDAGGRFEWAPVSI 42

RESULT 13
AAE25111
ID AAE25111 standard; protein; 222 AA.
XX
AC AAE25111;
XX
DT 30-OCT-2002 (first entry)
XX
DE Arabidopsis thaliana ICN7 full-length protein.
XX
KW Plant development; cyclin-dependent kinase; CDK inhibitor; ICN1; ICN2;
KW ICN2; ICN6; ICN7; ICN8; ICNK; morphogenesis; maturation; enlargement;
KW plant breeding; growth.
XX
OS Arabidopsis thaliana.
XX
PN WO200250292-A2.
XX
PD 27-JUN-2002.
XX
PF 18-DEC-2001; 2001WO-CA001825.
XX
PR 18-DEC-2000; 2000US-0255908P.
XX
PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.
PA (UYSA-) UNIV SASKATCHEWAN TECHNOLOGIES INC.
XX
PI Wang H, Zhou Y, Fowke LC;
XX WPI; 2002-519888/55.
DR N-PSDB; AAD40771.
XX
PT Modifying plant development, e.g. growth or maturation, which is
PT particularly useful in plant breeding, by introducing nucleic acids
PT coding for cyclin-dependent kinase (CDK) inhibitor, cyclin or proteins
PT that bind to CDK inhibitors.

PS Disclosure; Fig 10B; 89pp; English.
XX
CC The invention relates to a method for the development of a plant. The
CC method involves introducing into a plant cell a nucleic acid encoding a
CC protein that binds or interacts with a cyclin-dependent kinase (CDK)
CC inhibitor polypeptide (such as ICN1, ICN2, ICN6, ICN7, ICN8 and
CC ICNK), a cyclin polypeptide, a CDK, or a polypeptide that modulates the
CC degradation of a CDK inhibitor polypeptide. The method is useful for
CC modifying the growth and development of plants e.g. morphogenesis,
CC growth, multiplication, enlargement, differentiation or maturation of a
CC cell or plant. It is particularly useful in plant breeding. The present
CC sequence is A. thaliana ICN7 full-length protein

XX
SQ Sequence 222 AA;
Query Match 16.5%; Score 215.5; DB 5; Length 222;
Best Local Similarity 29.6%; Pred. No. 3.4e-11;
Matches 79; Conservative 34; Mismatches 93; Indels 61; Gaps 12;
QY 1 MGKYMRCRGAAGAEEVAEVTQW---GVRTRSPSAATGGVAKVAPRRKRAPAGEPA 56
Db 1 MGKYMKSK--ITGDISVMEVSKATAPSPGVRTRAKTLAL-----KRL---NSS 45
QY 57 AAVSAGDGGGSCYIHLRML-----FMAPPQPSVDSVPTVEAADGAAGQQAALAA 111

| | | | |
|----|-----|--|------------------|
| Db | 46 | AADSALPNDSSCYLQLRSRRLKPKPSSLIIEPKQP-PRVH | -----RSGIKESG 90 |
| QY | 112 | GLSRCSSTASSVNLGLGQGRSHTCRSYDAEAGGDHVLVDVSAASNS-GSGPDRERRET | 170 |
| Db | 91 | SRSRVDVNS-----APVAQSSNEDECDFNFVSQVSCGENSLGFESRHSHTRES | 139 |
| QY | 171 | TPSSRAHGLSDLESDLAGHKTPSLPAATP-----AAELIVPPAHEIOEFFFAAAEAAQAK | 226 |
| Db | 140 | TPCN-----FVEDMEIMVTPGSTRSMCRATKEYTREQDNVPTTSEMEEFFAYAEQQOQR | 195 |
| QY | 227 | RFASKYNFDFVRGVPLDAGGRFEWAPV | 253 |
| Db | 196 | LEMEKYNFNDIVNDIPL--SGRYEWVQV | 220 |

RESULT 14
AAB26252
ID AAB26252 standard; protein; 190 AA.
XX
AC AAB26252;
XX
DT 17-JAN-2001 (first entry)
XX
DE Corn cyclin-dependent kinase inhibitor #5.
XX
KW Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
KW cell growth; herbicide.
XX
OS Zea mays.
XX
PN WO200060087-A2.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US009106.
XX
PR 07-APR-1999; 99US-0128192P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Klein TM, Weng Z, Cahoon RE;
XX
DR WPI; 2000-679375/66.
DR N-PSDB; AAA95283.
XX

Cyclin dependent kinase inhibitor sequences, useful for identifying herbicides and plant growth inhibitors.
PS Claim 10; Fig 1; 58pp; English.
XX
CC The present sequence is the corn cyclin-dependent kinase inhibitor (CDKI). Its coding sequence was isolated by searching a contig of corn endosperm, leaf sheath, tassel and shoot cDNA for sequences similar to those encoding the CDKI from Chenopodium rubrum, Caenorhabditis elegans and Arabidopsis thaliana. CDKI is involved in the cell cycle, and may promote or inhibit cell division and growth. The protein and its coding sequence are useful in the production of transgenic plants which produce increased or decreased amounts of the CDKI protein, in the identification of herbicides, in genetic and physical mapping and in the isolation of the CDKI gene in other organisms
XX
SQ Sequence 190 AA;
Query Match 15.6%; Score 204; DB 3; Length 190;
Best Local Similarity 28.2%; Pred. No. 3e-10;
Matches 71; Conservative 34; Mismatches 79; Indels 68; Gaps 10;
QY 1 MGKYMRCRGAAGAEVAEVTQV-VGVRTRSRSAATGGVAKVAPRRKRAPAGEPAAV 59
Db 1 MGKYMRAK--ASSEVIMDVAAAPLGVTRARALAL-----QLQEQQQTQWEE 47
QY 60 SAGDGGGSCYIHLRSRMLFMAPPQPPQPSVDSVPTPVEADGAAGQQGAALAAGLSRCST 119

Db 48 GAGGE---YLELRNRRLKLP PPPP-----ATT 71
QY 120 ASSVNLGLGQGRSHTCRSYDAEAGGDHVLVDVSAASNSGSGPDRERRETTPSSRAHGE 179
Db 72 RRS-----GGRKAAAEAAATKEAEASYGENMLEAM-----ERITRETTPCSLINTQ 119
QY 180 LSDLE-SDLAGHKTPSLPAATPAAELIVPPAHEIOEFFFAAAEAAQAKRFASKYNFDFVR 238
Db 120 MTPGSTRSGHCHRRVNAAPPVHA---VPSREMNEYFAAEQRRQQQDFIDKYNFDPAN 176
QY 239 GVPLDAGGRFEW 250
Db 177 DCPLP--GRFEW 186
RESULT 15
AAB98862
ID AAB98862 standard; protein; 46 AA.
XX
AC AAB98862;
XX
DT 21-FEB-2003 (first entry)
XX
DE Maize CKI_D C-terminal protein fragment.
XX
KW Maize; cyclin-dependent kinase inhibitor; CKI; plant; corn; crop yield;
KW root size; plant growth; tassel size; ear size; male sterility;
KW endoreduplication.
XX
OS Zea mays.
XX
PN WO200281623-A2.
XX
PD 17-OCT-2002.
XX
PF 06-NOV-2001; 2001WO-US044038.
XX
PR 07-NOV-2000; 2000US-0246349P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
PA (ARIZ-) ARIZONA BOARD OF REGENTS.
XX
PI Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;
XX
DR WPI; 2003-058511/05.
XX
PT Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
PT involved in cell cycle regulation, and useful for altering cell cycle
PT protein content, cell cycle progression, cell number and composition of
PT plants.
XX
PS Example 11; Page 46; 69pp; English.
XX
CC The present invention relates to maize cyclin-dependent kinase inhibitor (CKI), CKI_B, CKI_C and CKI_D. The coding sequences (I) for these proteins are useful for modulating the activity of CDK in a plant such as a corn, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley, oil-seed Brassica and millet. Modulating the activity of CDK, preferably modulating downward is useful for providing differential growth in a plant, especially a positive growth advantage and modulating CDK activity upward is useful for increasing crop yield, root size, plant growth, tassel size and/or ear size. Modulating CDK activity is also useful for conferring male sterility and for improving transformation frequencies by increasing the number of cells in cell division. CDK activity can also be modulated for modulating endoreduplication in the endosperm of corn, sorghum, wheat, rice, barley, and millet, where the promoter used is an endosperm-preferred promoter. The cell numbers are modulated in one or more tissues of a plant, comprising root, seed, tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves, stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm, pericarp, meristematic or leaf cells. (I) is also useful for identifying maize CKI interacting proteins, by adducting the nucleic acid sequence to

CC a second nucleic acid sequence encoding a DNA-binding domain. The present
 CC sequence is a C-terminal protein fragment of CKI D. This sequence
 CC contains a CDK binding region and/or cyclin binding domains

XX

SQ Sequence 46 AA;

Query Match 15.5%; Score 202.5; DB 6; Length 46;
 Best Local Similarity 89.4%; Pred. No. 6.4e-11;
 Matches 42; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 210 AHEIQEFFFFAAQAQAKRFASKYNFDFVRGVPLDAGGRFEWAPVVS I 256
 Db 1 AQEIQEFFFFAAQAQAKRFASKYNFDFVRGVPLDA-GRFEWTPGVSI 46

Search completed: October 1, 2004, 16:51:15
 Job time : 123 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 16:49:11 ; Search time 32 Seconds
(without alignments)
413.008 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 1304
Sequence: 1 MGKYMRCRGAAGAEVAE.....VRGVPLDAGGRFEWAPVWSI 256

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 124 | 9.5 | 826 | 4 | US-09-894-998A-47 |
| 2 | 117.5 | 9.0 | 456 | 4 | US-09-252-991A-19417 |
| 3 | 112 | 8.6 | 477 | 4 | US-09-252-991A-24574 |
| 4 | 111 | 8.5 | 317 | 4 | US-09-252-991A-30984 |
| 5 | 111 | 8.5 | 425 | 4 | US-09-252-991A-31183 |
| 6 | 111 | 8.5 | 455 | 4 | US-09-252-991A-18699 |
| 7 | 110.5 | 8.5 | 1239 | 4 | US-09-252-991A-30198 |
| 8 | 109.5 | 8.4 | 605 | 4 | US-09-252-991A-27220 |
| 9 | 109 | 8.4 | 300 | 4 | US-09-252-991A-25073 |
| 10 | 108.5 | 8.3 | 726 | 4 | US-09-252-991A-20675 |
| 11 | 108.5 | 8.3 | 973 | 4 | US-09-252-991A-21386 |
| 12 | 108.5 | 8.3 | 1395 | 4 | US-09-252-991A-30345 |
| 13 | 108 | 8.3 | 362 | 4 | US-09-252-991A-17884 |
| 14 | 108 | 8.3 | 553 | 4 | US-09-252-991A-32621 |
| 15 | 107.5 | 8.2 | 534 | 4 | US-09-252-991A-22537 |
| 16 | 107.5 | 8.2 | 698 | 3 | US-08-941-445A-11 |
| 17 | 107 | 8.2 | 433 | 4 | US-09-252-991A-20044 |
| 18 | 106 | 8.1 | 560 | 4 | US-09-252-991A-27797 |
| 19 | 106 | 8.1 | 743 | 4 | US-09-252-991A-28327 |
| 20 | 106 | 8.1 | 871 | 4 | US-09-252-991A-19431 |
| 21 | 106 | 8.1 | 1046 | 4 | US-09-252-991A-16902 |
| 22 | 105.5 | 8.1 | 646 | 4 | US-09-252-991A-28206 |
| 23 | 104.5 | 8.0 | 478 | 4 | US-09-252-991A-19635 |
| 24 | 104.5 | 8.0 | 909 | 3 | US-09-425-383-2 |
| 25 | 104.5 | 8.0 | 2294 | 4 | US-09-252-991A-17231 |
| 26 | 104 | 8.0 | 395 | 4 | US-09-252-991A-17341 |
| 27 | 103 | 7.9 | 316 | 4 | US-09-252-991A-30877 |

| | | | | | | |
|----|-------|-----|------|---|----------------------|-------------------|
| 28 | 103 | 7.9 | 681 | 4 | US-09-490-291-4 | Sequence 4, Appli |
| 29 | 103 | 7.9 | 686 | 4 | US-09-252-991A-20509 | Sequence 20509, A |
| 30 | 103 | 7.9 | 691 | 4 | US-09-490-291-6 | Sequence 6, Appli |
| 31 | 102.5 | 7.9 | 445 | 4 | US-09-252-991A-23505 | Sequence 23505, A |
| 32 | 102.5 | 7.9 | 553 | 3 | US-09-083-351-2 | Sequence 2, Appli |
| 33 | 102.5 | 7.9 | 553 | 3 | US-09-083-352-2 | Sequence 2, Appli |
| 34 | 101 | 7.7 | 196 | 4 | US-09-252-991A-23071 | Sequence 23071, A |
| 35 | 101 | 7.7 | 214 | 4 | US-09-252-991A-31304 | Sequence 31304, A |
| 36 | 101 | 7.7 | 297 | 4 | US-09-252-991A-27909 | Sequence 27909, A |
| 37 | 101 | 7.7 | 335 | 2 | US-08-405-175A-6 | Sequence 6, Appli |
| 38 | 101 | 7.7 | 618 | 4 | US-09-252-991A-27666 | Sequence 27666, A |
| 39 | 100.5 | 7.7 | 347 | 4 | US-09-252-991A-19498 | Sequence 19498, A |
| 40 | 100.5 | 7.7 | 738 | 3 | US-08-864-038A-3 | Sequence 3, Appli |
| 41 | 100.5 | 7.7 | 931 | 4 | US-09-252-991A-22550 | Sequence 22550, A |
| 42 | 100.5 | 7.7 | 1093 | 3 | US-08-545-860D-55 | Sequence 55, Appl |
| 43 | 100.5 | 7.7 | 1093 | 5 | PCT-US94-04496-55 | Sequence 55, Appl |
| 44 | 100 | 7.7 | 311 | 4 | US-09-252-991A-18324 | Sequence 18324, A |
| 45 | 100 | 7.7 | 353 | 4 | US-09-252-991A-32769 | Sequence 32769, A |

ALIGNMENTS

RESULT 1
US-09-894-998A-47
; Sequence 47, Application US/09894998A
; Patent No. 6537555
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: Craig H. Day
; APPLICANT: Davin C. Dillon
; APPLICANT: McGowan, Patrick
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; FILE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538
; CURRENT APPLICATION NUMBER: US/09/894,998A
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 826
; TYPE: PRT
; ORGANISM: HSV-2
US-09-894-998A-47

| | | | | |
|-----------------------|--------|--|-------|-----------------------------------|
| Query Match | 9.5%; | Score 124; | DB 4; | Length 826; |
| Best Local Similarity | 28.3%; | Pred. No. 0.002; | | |
| Matches | 60; | Conservative | 18; | Mismatches 82; Indels 52; Gaps 9; |
| QY | 6 | RKCRGAAGAEVAAEVTVGVVTRSRSAATGGVAKVAPRRKRAPAGE-----PAAA | 58 | |
| Db | 514 | RKRRGSDSGPAASSASS--SAAPRSPLAPQGVGAKRAAPR--RAPDSGDRGHGFLAP | 569 | |
| QY | 59 | VSAGDGGSCYIHLRSRMLFMAPPQPPSVDSVPTPVEAADGAAGQQAALAAGLSRCSS | 118 | |
| Db | 570 | ASAGA-----APPSASPS-----SQAAVAAAASSSSASSSSSSSSAS | 606 | |
| QY | 119 | TASSVNLGLGGQSGSHTCRSYDAAGDGHVLDVSAASNSGSGPDRRETTPSSRAHG | 178 | |
| Db | 607 | SSASASSSSASSSSASSSSASSAGGAGG-----SVASASGAG---ERRETSGLPRAAA | 656 | |
| QY | 179 | ELSDLESDLAGHKT-----GPSLPAATPAEL | 205 | |
| Db | 657 | PRGPRK---CARKTRHAEGGPEFGARDPAPGL | 685 | |

RESULT 2
US-09-252-991A-19417
; Sequence 19417, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 19417
LENGTH: 456
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
FEATURE:
NAME/KEY: UNSURE
LOCATION: (62)
OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-19417

Query Match 9.0%; Score 117.5; DB 4; Length 456;
Best Local Similarity 29.9%; Pred. No. 0.0038;
Matches 66; Conservative 13; Mismatches 73; Indels 69; Gaps 13;
QY 26 GVTTRSRSAATGGVAKVAPRRKAPA-----GEP-----AAVSAGDGGSCYIHLR 73
Db GERRCAPASAGGAA--GKRRGPPARRQRTTEPATRGGGARPAAPGAAGS-----R 305
QY 74 SRMLFMAP-----PQPSVDSVPTVEAADGAAGQQAALAAAGL----- 113
Db GGARTGAPRPGRYNRPPIRRPRRPAQCAPVRHA-GAAAGGGAAGAGLRGGQLQGD 364
QY 114 -----SRCSTASSVNLGLG-----QGSHTCRSYDAEAGGDHVLVDVSAAS 157
Db ARTHPSRATGRAG-----GGHVPPPLARPRRQRGAGLRGIRAAAGQRRHQLHQRPA 418
QY 158 NSGS-GPDRRETTPSSRAHGELESDLAGH-KTGP 196
Db LPGAYPGRRRRRAGPA--ASGDVGDRH---RGHSRTGRSL 454

RESULT 3
US-09-252-991A-24574
Sequence 24574, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 24574
LENGTH: 477
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24574

Query Match 8.6%; Score 112; DB 4; Length 477;
Best Local Similarity 25.5%; Pred. No. 0.014;
Matches 70; Conservative 22; Mismatches 91; Indels 92; Gaps 14;
QY 4 YMRKCRGA-----AGAEVAAVEVTQVVGVRTRSRSAATGGVAKVAPRRK----- 48
Db WWRACRSASRWSPATSTTGASALINCSPVAACAKSSAPRAGRAASRRAGRCRWTAF 90
QY 49 --RAPAGEPAAAVSAGDGGSCYIHLR-----SRMLFMAPQPSVD-----SVTPV 95

Db 91 TCAMPAGASRAS-SRGGPGRCTCTTCRWRWRWSHELPLARRQPRRTADQRRRVLSAPVPV 149
QY 96 EAADGAAGQ-----QGAALAAG-----LSRCSSTASSVNL-----GLGGQRG 132
Db 150 HRR-GAPGDPGPNLHRLGRGRPAVAGSPQRRRRARCRGAGGGRRLHRLAQPLPGQAG 208
QY 133 -----SHTCRSYDAEAGGDHVLVDVSAASNSGSGPDRE---RRETT 171
Db 209 RLGRSRAPVRPAPAPARHAHQVSPVAPQAGD-----RSPAGLRRRHQLRRGPP 258
QY 172 PSSRAHGELESDLAGHKTGPS---LPAATPAA 203
Db 259 GTSRQHG-----QAGLRGARRGPGGARYPPGLPGA 288

RESULT 4
US-09-252-991A-30984
Sequence 30984, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 30984
LENGTH: 317
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30984

Query Match 8.5%; Score 111; DB 4; Length 317;
Best Local Similarity 25.3%; Pred. No. 0.0098;
Matches 73; Conservative 24; Mismatches 90; Indels 102; Gaps 13;
QY 13 GAEVAAVEVTQVVGVRTR-----SRSAATGG--VAKVAPRRKRAPAGEP---AAAV 59
Db 4 GAAMPGMAHTAVPRARYRVHARRYLLSRGSGAGGRWPAGVADRSAADPGLPLRRAAR 63
QY 60 SAGDGGSCYIHLRSMFLMAPPQPSVDSVPTPV-----EAADGAAGQ 105
Db 64 RAGGGAAC---LQRRPGVPAPHRQPAAGAAVYAVGRRLSRHGPFDDGLLEAGEGGAGHR 120
QY 106 GA-----ALAA-----GLSR-----CSSTASSVNLGLGGQSGS--- 133
Db 121 PAQGRDPFRLVATAAHRDADALRRRRRPPRPLPGPGCAERGETRLAAGRRRTGRQP 180
QY 134 ----HTCRSYDAEAGGDHVLVDVSAASNSGSGPDRETTTPSSRAHGELESDLAG 189
Db 181 LPRERSARGLPGRQAG-----LAAASAATGGAPRTL---LAAASAATGGAPRTL---LAR 215
QY 190 HKTGPSLPAATPAELIVP-----PAHEIQE-----FFAAAEAAQAKR 227
Db 216 GAGTAPAAQAPRAARTHPVAAGAPAAEEQDRPGGDRGHASAHRAPGRR 264

RESULT 5
US-09-252-991A-31183
Sequence 31183, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18

```

; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31183
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31183

```

Query Match 8.5%; Score 111; DB 4; Length 425;
Best Local Similarity 23.2%; Pred. No. 0.015;
Matches 76; Conservative 24; Mismatches 104; Indels 124; Gaps 14;

| | | | |
|----|-----|---|-----|
| QY | 9 | RGAAGA EV-AAVEVTQVVGVTRTSAAAATGGVAKVAPRRKRPAFAGEPAAAAVSAGGDG- | 66 |
| Db | 106 | RCAAGAAAGAAATELPAALGLGYRSQPVA---GAAEA-----QLPGGPPVAAVDPEGRRPGP | 157 |
| QY | 67 | -----SCYTHLSRMLFMA-----PPOFPQSVDSV----- | 91 |
| Db | 158 | ARPVRGORRLATACLHRA--MADRRDDEPDRRAGDPVLHPFRSRPSAAGADRLAQ | 213 |
| QY | 92 | --PTPVEAADGAAGQGGAALAAGLSRC-----SSTASSVN----- | 125 |
| Db | 214 | WHPAGLRAG-GAAHRGRGHDAADALACRTVRRPRWCATVPVALARSHLAGAAAALVSAP | 272 |
| QY | 126 | -----GLGGQRGSHTCRSYDAAEAGGDHVLYDVSAASNSGSGPDRE | 166 |
| Db | 273 | GIFLDRRGAAWHGALPEPGGAEPGA-----ARAGGGRHGVRDPGDAGTGHSRPYRS | 325 |
| QY | 167 | RRETTTPSSRAHGE LSDLES DL AGHKT-----GPSLPATPAABELIVPPAHEIOEFFAAA | 220 |
| Db | 326 | RAATA-----GSDALGFRSAQSR LR GAGVPAQP VAGELGAAPGRRRTVGFGVPA | 373 |
| QY | 221 | E----AAQAKRFASKYNFD FVRGVPLDA | 244 |
| Db | 374 | VRLVLCADAVSSAGRRPSPRLIRGLAADA | 401 |

RESULT 6
US-09-252-991A-18699
; Sequence 18699, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNO
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18699
; LENGTH: 455
; TYPE: prt
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18699

Query Match 8.5%; Score 111; DB 4; Length 455;
 Best Local Similarity 24.9%; Pred. No. 0.016;
 Matches 70; Conservative 26; Mismatches 127; Indels 58; Gaps 11;

| | | | |
|----------------|-----|---|-----|
| Qy | 8 | CRGAAGAEVAAVEVTQVGVRTSRSAATGGVAKVAPRRKRPAAGEP-----AA | 57 |
| D _b | 63 | CRSAA-LDAPGTSLQLPARSRTLQPLSGGGRPAWGDRGFQPAGVFQGRHLRAHAGAR | 121 |
| Qy | 58 | AVSAGGDGGSC-----YIHLRSRM----LFMAPPQPQPSVDSPVPVEAADGAAG---- | 103 |
| D _b | 122 | VHQAGPGGRTARADRLDHGQQQRLSAAGAAEADGHAAVAVROOP-GVRAAGGGAAPDPAPV | 180 |

```

QY 104 ---QQGAALAAGLRCSSTASSVNLGLGGQSGS---HTCRSY-----DAAEAGGDHVL- 150
Db 181 LQLPRGAACRLGTERAGAAVPAVATQPGVLAANPRHLCQRFQPEGVPDLLAAVVEHQLF 240
QY 151 -----VDVSAASNSGSGPDRERRETTTPSSRAHGE LSDL EAGKGTGSP LPAATPAE L 205
Db 241 AGNPVAEDPPAGDERRGNARLRRRPRPSGERSGFLQVHREHVAGR-----AVQH 289
QY 206 IVPPAHEIQEFFAAEAQAQRFASKYN---FDFVRGVPLD 243
Db 290 VVGSGETEQQRQAVATVAADHDQVASLFRGEAVDFLARLPVD 330

```

RESULT 7
US-09-252-991A-30198
; Sequence 30198, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND A
; TITLE OF INVENTION: AERUGINOSA FOR DIA
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,9
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,78
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,19
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30198
; LENGTH: 1239
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30198

Query Match 8.5%; Score 110.5; DB 4; Length 1239;
Best Local Similarity 23.2%; Pred. NO. 0.069;
Matches 59; Conservative 25; Mismatches 93; Indels 77; Gaps 10;

| | | |
|----|-----|---|
| QY | 8 | CRGAAGAEVAAVE--VTQVVGVTRTSRAAATGGVAKVAPRRKRAPAGEPAA-----57 |
| | | : : : : : : : : : : |
| Db | 141 | CRAAAEALDAVDKPVQALRGITLAAAEQLVAQGGPGRLAPEIEREAVGLGEALGGCHDQR 200 |
| QY | 58 | -----AVSAGGDG-----GSCYIHLR-----SRMLFMAPPQPQPSVDS 90 |
| | | : : : : : : : |
| Db | 201 | GRIGORHETDVQALFRFCVGTAGPGLAAYRSVCHHLKCGRAARAVPFSSAPAPSSARFS 260 |
| QY | 91 | VPTPVEA----ADGAAGCQQAALAAGL-----SRCSSSTASSVNLGLGG 129 |
| | | : : : : : : |
| Db | 261 | IAATASAMPGGAGGFAGQEKAPHGAGLVRRRVAGSGLDDRDGLVRSRFAELGDLVEAQVAV 320 |
| QY | 130 | QRGSHTCRSYDAAEAGGDHVLV-----DVSAANSKSGSPDRERRETTPSSRA 176 |
| | | : : : : : : : |
| Db | 321 | DVGSH---AVDVLRVG--HVLLHGRQFAEDRGDRRRQLRCQEHIGGIAQTVREVTGRSGD 375 |
| QY | 177 | HGELSDES DLAGH 190 |
| | | : : : |
| Db | 376 | HGG-TGLDASLVAH 388 |

RESULT 8
US-09-252-991A-27220
; Sequence 27220, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AL
; TITLE OF INVENTION: AERUGINOSA FOR DIA
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,9
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,78

Job time : 34 secs

Query Match 8.3%; Score 108; DB 4; Length 553;

| | | | |
|----|-----|---|-----|
| QY | 11 | AAGAEVAAVEVTQVGVTRTRSRSAATCGVAKVAPRRKRAPAGEPAAVASAGDGGSCYI | 70 |
| Db | 302 | AAHAALAAPDRSVPGPAADRSPAAGAH-----ARRAKRAPA-RPAAAVASGG----- | 347 |
| QY | 71 | HLRSRMLEMAPPPQPOPSVDSVPTPVEAADGAAGQQAALAGLSRCSSTASSVNLGLGGQ | 130 |
| Db | 348 | ---GLDLALA---GQPLACRSGAPAAALGGSPGR---VAAALRRVAPRAQPVQ-----Q | 392 |
| QY | 131 | RGS--HTC-----RSYDAAAEAGGDHVLVDVSAASNSGSGPDRRRETTTPSSRAHGEL | 180 |
| Db | 393 | RGAAHRLCPRCAARLRPAKDAAAGGRSH-----TGTRAGAAERRTRTGPGEIATGDC | 444 |
| QY | 181 | SDLESDLAGHKTGPSLPAATPAAEELIVPPAHEIQEFFFFAAEAAQAQKRFASKYNFDVVRGV | 240 |
| Db | 445 | QOPAAE-----ETVPAACPAAAHRPEQPG----- | 467 |

RESULT 15
US-09-252-991A-22537
; Sequence 22537, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22537
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22537

| | | | | |
|-----------------------|--------|-------------------|-----------------------|---------------------------------------|
| Query Match | 8.2%; | Score 107.5; | DB 4; | Length 534; |
| Best Local Similarity | 27.2%; | Pred. No. 0.043; | | |
| Matches | 58; | Conservative | 18; | Mismatches 78; Indels 59; Gaps 12; |
| QY | 9 | RGAGAEVAAVEVTQVGV | RRRSAAATGGVAKVAPRRKRA | -----PAGEPAAAVSAGG 63 |
| Db | 127 | RGTA-VQPRLV | ALRAVAEPTAGRG | AHRTG-----RRRRDADAGGTAGDP-ARLPAPG 177 |
| QY | 64 | DGSCYIHLRSRMLFM | APPQPPQPSVDSP | TPVEADGAAGCQQAALAAGLSRCSSSTASSV 123 |
| Db | 178 | PGACAQLRHHSR | -----PARP | -----AGGTACGLAGGVRLIA----- 210 |
| QY | 124 | NLGLGGQSGSHTCR | --SYDAAEAGGDHVLVDV | SAANSNGSGPDRERRETTPS-SRAHGEI 180 |
| Db | 211 | -LDLPDQPAGRPARLP | GGDEADAG----- | PAQPGAEPFRQHR--LPALRRILHG-- 255 |
| QY | 181 | SDLESDLAGHKTGPS | LPAATPAAE | LIVPPAHEI 213 |
| Db | 256 | ADLHRPGRTRRAAP | VAAPARGPAADRET | GPAPHL 288 |

Search completed: October 1, 2004, 16:55:10

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 16:53:56 ; Search time 130 Seconds
(without alignments)
633.697 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 1304
Sequence: 1 MGKYMRCRGAAGAEVAE.....VRGVPLDAGRFEPWAPVWSI 256

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|----------------------|
| 1 | 1304 | 100.0 | 256 | 9 | US-09-993-308-2 |
| 2 | 1304 | 100.0 | 256 | 11 | US-09-993-808B-2 |
| 3 | 1217.5 | 93.4 | 263 | 12 | US-10-425-114-63204 |
| 4 | 1059.5 | 81.2 | 221 | 12 | US-10-425-114-69245 |
| 5 | 669 | 51.3 | 262 | 15 | US-10-333-006-10 |
| 6 | 669 | 51.3 | 417 | 16 | US-10-437-963-195115 |
| 7 | 668.5 | 51.3 | 278 | 12 | US-10-425-114-71211 |
| 8 | 665 | 51.0 | 227 | 9 | US-09-993-308-6 |
| 9 | 665 | 51.0 | 227 | 11 | US-09-993-808B-6 |
| 10 | 399.5 | 30.6 | 116 | 15 | US-10-333-006-14 |
| 11 | 315.5 | 24.2 | 354 | 16 | US-10-437-963-122432 |
| 12 | 314 | 24.1 | 423 | 16 | US-10-437-963-122443 |
| 13 | 279 | 21.4 | 226 | 15 | US-10-333-006-55 |
| 14 | 238 | 18.3 | 194 | 16 | US-10-437-963-120870 |
| 15 | 235 | 18.0 | 194 | 15 | US-10-333-006-44 |

| | | | | | | |
|----|-------|------|-----|----|----------------------|-------------------|
| 16 | 231 | 17.7 | 255 | 12 | US-10-425-114-61054 | Sequence 61054, A |
| 17 | 226 | 17.3 | 218 | 16 | US-10-437-963-128205 | Sequence 128205, |
| 18 | 220.5 | 16.9 | 225 | 16 | US-10-437-963-198574 | Sequence 198574, |
| 19 | 218 | 16.7 | 205 | 12 | US-10-424-599-235800 | Sequence 235800, |
| 20 | 215.5 | 16.5 | 222 | 16 | US-10-451-139-21 | Sequence 21, Appl |
| 21 | 215 | 16.5 | 248 | 12 | US-10-425-114-59718 | Sequence 59718, A |
| 22 | 212.5 | 16.3 | 222 | 12 | US-10-688-291-4 | Sequence 4, Appli |
| 23 | 205.5 | 15.8 | 224 | 12 | US-10-424-599-210190 | Sequence 210190, |
| 24 | 200 | 15.3 | 190 | 9 | US-09-993-308-4 | Sequence 4, Appli |
| 25 | 200 | 15.3 | 190 | 11 | US-09-993-808B-4 | Sequence 4, Appli |
| 26 | 178 | 13.7 | 196 | 9 | US-09-733-507-16 | Sequence 16, Appl |
| 27 | 178 | 13.7 | 196 | 16 | US-10-451-139-15 | Sequence 15, Appl |
| 28 | 178 | 13.7 | 289 | 16 | US-10-451-139-17 | Sequence 17, Appl |
| 29 | 160 | 12.3 | 205 | 12 | US-10-424-599-182928 | Sequence 182928, |
| 30 | 153 | 11.7 | 85 | 16 | US-10-767-701-37255 | Sequence 37255, A |
| 31 | 150.5 | 11.5 | 176 | 12 | US-10-424-599-212181 | Sequence 212181, |
| 32 | 149 | 11.4 | 327 | 12 | US-10-688-291-34 | Sequence 34, Appl |
| 33 | 145 | 11.1 | 137 | 16 | US-10-451-139-13 | Sequence 13, Appl |
| 34 | 144.5 | 11.1 | 108 | 15 | US-10-333-006-16 | Sequence 16, Appl |
| 35 | 142.5 | 10.9 | 137 | 9 | US-09-733-507-14 | Sequence 14, Appl |
| 36 | 134 | 10.3 | 90 | 15 | US-10-333-006-13 | Sequence 13, Appl |
| 37 | 134 | 10.3 | 196 | 16 | US-10-451-139-22 | Sequence 22, Appl |
| 38 | 134 | 10.3 | 212 | 9 | US-09-733-507-12 | Sequence 12, Appl |
| 39 | 134 | 10.3 | 212 | 16 | US-10-451-139-11 | Sequence 11, Appl |
| 40 | 132.5 | 10.2 | 180 | 12 | US-10-425-114-49866 | Sequence 49866, A |
| 41 | 132 | 10.1 | 95 | 16 | US-10-767-701-36263 | Sequence 36263, A |
| 42 | 129.5 | 9.9 | 93 | 15 | US-10-333-006-15 | Sequence 15, Appl |
| 43 | 129.5 | 9.9 | 204 | 16 | US-10-437-963-169561 | Sequence 169561, |
| 44 | 127.5 | 9.8 | 175 | 12 | US-10-425-114-46534 | Sequence 46534, A |
| 45 | 127.5 | 9.8 | 176 | 12 | US-10-425-114-61692 | Sequence 61692, A |

ALIGNMENTS

RESULT 1
US-09-993-308-2
; Sequence 2, Application US/099933308
; Patent No. US20020159435A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William J.
; APPLICANT: Lowe, Keith S.
; APPLICANT: Larkins, Brian A.
; APPLICANT: Dilkes, Brian R.
; APPLICANT: Sun, Yuejin
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,308
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 60/246,349
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Zea mays
US-09-993-308-2

| | | | | |
|-----------------------|-----------------|---|-----------|-------------|
| Query Match | 100.0% | Score 1304; | DB 9; | Length 256; |
| Best Local Similarity | 100.0% | Pred. No. 1e-101; | | |
| Matches 256; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| QY | 1 | MGKYMRCRGAAGAEVAEVTQVGVTRTSRRAATGGVAKVAPRRKRAPAGEPAAAVS | 60 | |
| Db | 1 | MGKYMRCRGAAGAEVAEVTQVGVTRTSRRAATGGVAKVAPRRKRAPAGEPAAAVS | 60 | |
| QY | 61 | AGDGGGSCYIHLRSMLFMAPPQPOPSVDSVPTVEAADGAAGQQAALAGLRCSSTA | 120 | |
| Db | 61 | AGDGGGSCYIHLRSMLFMAPPQPOPSVDSVPTVEAADGAAGQQAALAGLRCSSTA | 120 | |
| QY | 121 | SSVNLGLGGQSGHTCRSYDAEAGGDHVLVDVSAANSNGSGPDRETTTSSRAHGEL | 180 | |
| Db | 121 | SSVNLGLGGQSGHTCRSYDAEAGGDHVLVDVSAANSNGSGPDRETTTSSRAHGEL | 180 | |

Db 121 SSVNLGLGGQSGSHTCRSYDAEAGGDHVLVDVSAASNSGSGPDRRERTTSSRAHGE 180
QY 181 SDLESDLAGHKTPSLPAATPAEELIVPPAHEIQEFFFFAAAEAAQAKRFASKYNFDFVRGV 240
Db 181 SDLESDLAGHKTPSLPAATPAEELIVPPAHEIQEFFFFAAAEAAQAKRFASKYNFDFVRGV 240
QY 241 PLDAGGRFEWAPVVS 256
Db 241 PLDAGGRFEWAPVVS 256

RESULT 2

US-09-993-808B-2
; Sequence 2, Application US/09993808B
; Publication No. US20040003433A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William
; APPLICANT: Lowe, Keith
; APPLICANT: Sun, Yuejin
; APPLICANT: Dilkes, Brian
; APPLICANT: Larkins, Brian
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides,
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,808B
; CURRENT FILING DATE: 2001-11-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 256
; TYPE: PRT
; ORGANISM: zea mays
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(256)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-993-808B-2

Query Match 100.0%; Score 1304; DB 11; Length 256;
Best Local Similarity 100.0%; Pred. No. 1e-101;
Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGKYMRCRGAAGAEVAAVEVTQVVGVRTRSRSAATGGVAKVAPRRKRAPAGEPAAVVS 60
Db 1 MGKYMRCRGAAGAEVAAVEVTQVVGVRTRSRSAATGGVAKVAPRRKRAPAGEPAAVVS 60
QY 61 AGDGGGSCYIHLRSMLFMAPPQPOPSVDSVPTPVEAADGAGQGAALAAGLSRCSSTA 120
Db 61 AGDGGGSCYIHLRSMLFMAPPQPOPSVDSVPTPVEAADGAGQGAALAAGLSRCSSTA 120
QY 121 SSVNLGLGGQSGSHTCRSYDAEAGGDHVLVDVSAASNSGSGPDRRERTTSSRAHGE 180
Db 121 SSVNLGLGGQSGSHTCRSYDAEAGGDHVLVDVSAASNSGSGPDRRERTTSSRAHGE 180
QY 181 SDLESDLAGHKTPSLPAATPAEELIVPPAHEIQEFFFFAAAEAAQAKRFASKYNFDFVRGV 240
Db 181 SDLESDLAGHKTPSLPAATPAEELIVPPAHEIQEFFFFAAAEAAQAKRFASKYNFDFVRGV 240
QY 241 PLDAGGRFEWAPVVS 256
Db 241 PLDAGGRFEWAPVVS 256

RESULT 3

US-10-425-114-63204
; Sequence 63204, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 63204
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17355F05_FLI.pap
US-10-425-114-63204

Query Match 93.4%; Score 1217.5; DB 12; Length 263;
Best Local Similarity 95.3%; Pred. No. 2.1e-94;
Matches 245; Conservative 3; Mismatches 4; Indels 5; Gaps 3;

QY 1 MGKYMRCRGAAGAEVAAVEVTQVVGVRTRSRSAATGGVAK-VAPRRKRAPAGEPAAVVS 59
Db 11 MGKYMRCRGAAGAEVAAVEVTQVVGVRTRSRSAATGGVAKVAPRRKRAPAGEPAAVVS 70
QY 60 SAGDGGGSCYIHLRSMLFMAPPQPOPSVDSVPTPVEAADGAGQGAALAAGLSRCSSST 119
Db 71 GAGDGGGSCYIHLRSMLFMAPPQPO--PSVTPFAAADGAGQGAALAAGLSRCSSST 128
QY 120 ASSVNLGLGGQSGSHTCRSYDAEAGGDHVLVDVSAASNSGSGPDRRERTTSSRAHGE 179
Db 129 ASSVH--VGGQSGSHTCRSDDAAEAGGDHVLVDVSAASNSGSGPDRRERTTSSRAHGE 186
QY 180 LSDLESDLAGHKTPSLPAATPAEELIVPPAHEIQEFFFFAAAEAAQAKRFASKYNFDFVRG 239
Db 187 LSDLESDLAGHKTPSLPAATPAEELIVPPAHEIQEFFFFAAAEAAQAKRFASKYNFDFVRG 246
QY 240 VPLDAGGRFEWAPVVS 256
Db 247 VPLDAGGRFEWAPVVS 263

RESULT 4

US-10-425-114-69245
; Sequence 69245, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 69245
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3079-036-F9_FLI.pap
US-10-425-114-69245

Query Match 81.2%; Score 1059.5; DB 12; Length 221;
Best Local Similarity 94.7%; Pred. No. 3.4e-81;
Matches 213; Conservative 3; Mismatches 4; Indels 5; Gaps 3;

QY 33 SAAATGGVAK-VAPRRKRAPAGEPAAVVSAGDGGGSCYIHLRSMLFMAPPQPOPSVDSV 91
Db 1 SAAATGGVAKVAPRRKRAPAGEPAAVVSAGDGGGSCYIHLRSMLFMAPPQPO--PSV 58

QY 92 PTPVEAADGAGCGGAAAGLRCSTASSVNLGLGGQSGSHTCRSYDAEAGGDHVLV 151
Db 59 PTPAEAADGAGCGGAAAGLRCSTASSVH--VGGQSGSHTCRSDDAEAGGDHVLV 116
QY 152 DVSAASNSGSGPDRERRETTSPSSRAHGELSDLESDLAGHKTGPSLPAATPAAELIVPPAH 211
Db 117 DVSAASNSGSGPDRERRETTSPSSRAHGELSDLESDLAGHKTGPSLPAATPAAELIVPPAH 176
QY 212 EIQEFFAAAEAAQAARFASKYNFDFVRGVPLDAGGRFEWAPVVS 256
Db 177 EIQEFFAAAEAAQAARFASKYNFDFVRGVPLDAGGRFEWAPVVS 221

RESULT 5
US-10-333-006-10
; Sequence 10, Application US/10333006
; Publication No. US20040019926A1
; GENERAL INFORMATION:
; APPLICANT: Frankard, Valerie Marie-No. US20040019926A1lle S.
; APPLICANT: Peres Bota, Adrian Marius
; APPLICANT: Droual, Anne-Marie
; APPLICANT: Mironov, Vladimir
; APPLICANT: Inz., Dirk
; APPLICANT: Hatzfeld, Yves
; TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
; FILE REFERENCE: 1187-13
; CURRENT APPLICATION NUMBER: US/10/333,006
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: PCT/IB01/01492
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/218,471
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/241,219
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-333-006-10

Query Match 51.3%; Score 669; DB 15; Length 262;
Best Local Similarity 60.3%; Pred. No. 3.2e-48;
Matches 161; Conservative 19; Mismatches 67; Indels 20; Gaps 10;

QY 1 MGKYMRCRGAAGAEVAAVEVTQVVGVRTRSRSAATGG-----VAKVAPRRKRAPAGE 54
Db 1 MGKYMRCRFGATGEELAAAMEVTQVVGVRTRSRSAAGATTTKVQAASAASTRRRKALLP 60
QY 55 PAAAVSAGDGGGCGYIHLRSMLFMAPPQPSVDSVPTPVEAADGAGCGGAAAGL 114
Db 61 TAVVGTTRRDGGGCGYQLRSMLFMAPPAPAA-RAPVVAEAA-GSGNGAAHAAAGLS 118
QY 115 RCSSTASSVNLGLGGQSGSHTCRSYDAEAGGDHVLVDVSAASNSGSGPDRERRETTSPSS 174
Db 119 RCSSTASSVD--AAQDRSLACRS-DVAEAGSEH--VPEGSASDSASGRDRERRETTSPSS 173
QY 175 RAHGELSDLESDLA-GHKTGPSLP-AATPAAELI----VPPAHEIQEFFAAAEAAQAARF 228
Db 174 FLPGEVSDLESDLAGGQKRSRPLPSAATASAQQATRPKIPPAEIEAFFAAAEAEAKRF 233
QY 229 ASKYNFDFVRGVPLDAGGRFEWAPVVS 255
Db 234 AAKYNFDFVRGVPLDA-GRFEWTPVVS 259

RESULT 6
US-10-437-963-195115
; Sequence 195115, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 195115
; LENGTH: 417
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91095C.1.1.pap
US-10-437-963-195115

Query Match 51.3%; Score 669; DB 16; Length 417;
Best Local Similarity 60.3%; Pred. No. 5.6e-48;
Matches 161; Conservative 19; Mismatches 67; Indels 20; Gaps 10;

QY 1 MGKYMRCRGAAGAEVAAVEVTQVVGVRTRSRSAATGG-----VAKVAPRRKRAPAGE 54
Db 156 MGKYMRCRFGATGEELAAAMEVTQVVGVRTRSRSAAGATTTKVQAASAASTRRRKALLP 215
QY 55 PAAAVSAGDGGGCGYIHLRSMLFMAPPQPSVDSVPTPVEAADGAGCGGAAAGL 114
Db 216 TAVVGTTRRDGGGCGYQLRSMLFMAPPAPAA-RAPVVAEAA-GSGNGAAHAAAGLS 273
QY 115 RCSSTASSVNLGLGGQSGSHTCRSYDAEAGGDHVLVDVSAASNSGSGPDRERRETTSPSS 174
Db 274 RCSSTASSVD--AAQDRSLACRS-DVAEAGSEH--VPEGSASDSASGRDRERRETTSPSS 328
QY 175 RAHGELSDLESDLA-GHKTGPSLP-AATPAAELI----VPPAHEIQEFFAAAEAAQAARF 228
Db 329 FLPGEVSDLESDLAGGQKRSRPLPSAATASAQQATRPKIPPAEIEAFFAAAEAEAKRF 388
QY 229 ASKYNFDFVRGVPLDAGGRFEWAPVVS 255
Db 389 AAKYNFDFVRGVPLDA-GRFEWTPVVS 414

RESULT 7
US-10-425-114-71211
; Sequence 71211, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 71211
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3078-022-G5_FLI.pap
US-10-425-114-71211
Query Match 51.3%; Score 668.5; DB 12; Length 278;
Best Local Similarity 73.1%; Pred. No. 3.8e-48;

Matches 174; Conservative 7; Mismatches 30; Indels 27; Gaps 17;
QY 1 MGKYMRCRGAAGAEVAAVEVTQVGVTRSRSAAT-GGVAKVA-PRRKRA--PA---- 52
Db 47 MGKYMRCRGAAGAEVAAVEVSQVGVTRSRSAATGGVAKVAPPRKKALLPAANET 106
QY 53 --GEPAAAVSAGDGGG-CYIHLRSRMLFMAPPQPSVDSVTPVVEAADGAAGQOG-AA 108
Db 107 ASGEPGAVGAGGCGGCCCYIHLRSRMLFMAPPQPS--AALTPVEAA-GAAXQGGVVA 163
QY 109 LAAGLSRCSSTASSVNLGLGGQR---GSHTCRSYDA-AEAGGDHVLVDVSAASNSGSGPD 164
Db 164 LAAGLSRCSSTASSVD--VGGQPPASGSHACRSDAAPAEVDGDHV-PDVVTASNSGSGVPD 220
QY 165 RERETTP-SSRAH-GELSDLESDLAG-HKTG-PSLPA-ATPAAELIVPPAHEIQEFF 217
Db 221 RERETTPSSRAHGGELSDLESDLVGRQKTGCSSSPATTTSAELIVPPAHEIQEFF 278

RESULT 8

US-09-993-308-6
; Sequence 6, Application US/09993308
; Patent No. US20020159435A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William J.
; APPLICANT: Lowe, Keith S.
; APPLICANT: Larkins, Brian A.
; APPLICANT: Dilkes, Brian R.
; APPLICANT: Sun, Yuejin
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides and Uses Thereof.
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,308
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 60/246,349
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(111)
; OTHER INFORMATION: The 'Xaa' at location 111 stands for Gln.
US-09-993-308-6

Query Match 51.0%; Score 665; DB 9; Length 227;
Best Local Similarity 72.6%; Pred. No. 5.8e-48;
Matches 172; Conservative 6; Mismatches 31; Indels 28; Gaps 16;
QY 1 MGKYMRCRGAAGAEVAAVEVTQVGVTRSRSAAT-GGVAKVA-PRRKRA--PA---- 52
Db 1 MGKYMRCRGAAGAEVAAVEVSQVGVTRSRSAATGGVAKVAPPRKKALLPAANVT 60
QY 53 --GEPAAAVSAGDGGG-CYIHLRSRMLFMAPPQPSVDSVTPVVEAADGAAGQOG-AA 108
Db 61 TSGEPAVGAGGCGGCCCYIHLRSRMLFMAPPQPS--AALTPVEAA-GAAXQGGVVA 117
QY 109 LAAGLSRCSSTASSVNLGLGGQRGSHTCRSYDA-AEAGGDHVLVDVSAASNSGSGPD 167
Db 118 LAAGLSRCSSTASSVDV-----GGHACRSDAAPAEVDGDHV-PDVVTASNSGSGVPDR 170
QY 168 RETTP-SSRAH-GELSDLESDLAG-HKTG-PSLPA-ATPAAELIVPPAHEIQEFFAA 219
Db 171 RETTPSSRAHGGELSDLESDLVGRQKTGCSSSPATTTSAELIVPPAHEIQEFFAA 227

RESULT 9

US-09-993-808B-6
; Sequence 6, Application US/09993808B
; Publication No. US2004003433A1
; GENERAL INFORMATION:

; APPLICANT: Gordon-Kamm, William
; APPLICANT: Lowe, Keith
; APPLICANT: Sun, Yuejin
; APPLICANT: Dilkes, Brian
; APPLICANT: Larkins, Brian
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides,
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,808B
; CURRENT FILING DATE: 2001-11-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 227
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(227)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-993-808B-6

Query Match 51.0%; Score 665; DB 11; Length 227;
Best Local Similarity 72.6%; Pred. No. 5.8e-48;
Matches 172; Conservative 6; Mismatches 31; Indels 28; Gaps 16;
QY 1 MGKYMRCRGAAGAEVAAVEVTQVGVTRSRSAAT-GGVAKVA-PRRKRA--PA---- 52
Db 1 MGKYMRCRGAAGAEVAAVEVSQVGVTRSRSAATGGVAKVAPPRKKALLPAANVT 60
QY 53 --GEPAAAVSAGDGGG-CYIHLRSRMLFMAPPQPSVDSVTPVVEAADGAAGQOG-AA 108
Db 61 TSGEPAVGAGGCGGCCCYIHLRSRMLFMAPPQPS--AALTPVEAA-GAAXQGGVVA 117
QY 109 LAAGLSRCSSTASSVNLGLGGQRGSHTCRSYDA-AEAGGDHVLVDVSAASNSGSGPD 167
Db 118 LAAGLSRCSSTASSVDV-----GGHACRSDAAPAEVDGDHV-PDVVTASNSGSGVPDR 170
QY 168 RETTP-SSRAH-GELSDLESDLAG-HKTG-PSLPA-ATPAAELIVPPAHEIQEFFAA 219
Db 171 RETTPSSRAHGGELSDLESDLVGRQKTGCSSSPATTTSAELIVPPAHEIQEFFAA 227

RESULT 10

US-10-333-006-14
; Sequence 14, Application US/10333006
; Publication No. US20040019926A1
; GENERAL INFORMATION:
; APPLICANT: Frankard, Valerie Marie-No. US20040019926A1lle S.
; APPLICANT: Peres Bota, Adrian Marius
; APPLICANT: Droual, Anne-Marie
; APPLICANT: Mironov, Vladimir
; APPLICANT: Inz, Dirk
; APPLICANT: Hatzfeld, Yves
; TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
; FILE REFERENCE: 1187-13
; CURRENT APPLICATION NUMBER: US/10/333,006
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: PCT/IB01/01492
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/218,471
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/241,219
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(11)

; OTHER INFORMATION: Xaa = any amino acid
US-10-333-006-14

Query Match 30.6%; Score 399.5; DB 15; Length 116;
Best Local Similarity 81.9%; Pred. No. 5.8e-26;
Matches 95; Conservative 0; Mismatches 14; Indels 7; Gaps 7;

QY 146 GDHVLVDVSAASNSGSGPDRERETTP-SSRAH-GELSDLESDLAG-HKTG-PSLPA-AT 200
Db 3 GDHV-PDVVXASNSGSGPDRERETTPSSRAHGGELSDLESDLVGRQKTCSSSPATTT 61
QY 201 PAELIVPPAHEIOEFFAAAEAAQAKRFASKYNFDFVRGVPLDAGGRFEWAPVVS 256
Db 62 SAELIVPPAHEIOEFFAAAEAAHAKRFASKYNFDFVRGVPLDA-GRFEWTPGVSI 116

RESULT 11

US-10-437-963-122432
; Sequence 122432, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 122432

; LENGTH: 354
; TYPE: PRT

; ORGANISM: Oryza sativa
; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_2535C.1.pep
US-10-437-963-122432

Query Match 24.2%; Score 315.5; DB 16; Length 354;
Best Local Similarity 37.4%; Pred. No. 2.7e-18;
Matches 105; Conservative 26; Mismatches 81; Indels 69; Gaps 13;

QY 1 MGKMRKCRGAAGAEVAAVEVTQVGVTRRSRAATGGVAKVAPRRKRAPAGEPAAAVS 60
Db 1 MGK-KKKRDGAARRQARVVVG-----AAVTARRVVASAEEGCGLVG 46
QY 61 AGDGGG-----CYIHLRSRLFMAPPQPPSVDSVPTPVEAADGAAGQQAALAAAG 112
Db 47 RGGGGGGGDDGEGGCYLRRL-----PFVAAAVVSSRREALGDS 90
QY 113 LSRCSSTASSVNLGQGRGSHTCRSYD-AAEAGGDHVLVDVSAASNSGSGPDRERETT 171
Db 91 VAEAAASSSSRAVELLGCSEEEAAEKVCTQAGEDH--DEESSVDSGCG--RERSATT 146
QY 172 PSSR---AHGELSLESD-----LAGHKTGSPSLPAATPAELIVPP--AH 211
Db 147 PSSRRPPGDADSSDAESNQEAKQCMCRSSSTTSAAPHAG---ATRSFRMAPPAAAA 202
QY 212 EIOEFFAAAEAAQAKRFASKYNFDFVRGVPLDAG--GRFEW 250
Db 203 EIEEFLAAERSEAEERFAAKYNFDFVRGVPLDAGGAGRFEW 243

RESULT 12

US-10-437-963-122443
; Sequence 122443, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 122443
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_2536C.1.pep
US-10-437-963-122443

Query Match 24.1%; Score 314; DB 16; Length 423;
Best Local Similarity 36.4%; Pred. No. 4.5e-18;
Matches 107; Conservative 27; Mismatches 78; Indels 82; Gaps 14;

QY 1 MGKMRKCRGAAGAEVAAVEVTQVGVTRRSRAATGGVAKVAPRRKRAPAGEPAAAVS 60
Db 1 MGK-KKKRDGAARRQARVVVG-----AAVTARRVVASAEEGCGLVG 46
QY 61 AGDGGG-----CYIHLRSRLFMAPPQPPSVDSVPTPVEAADGAAGQQAALAAAG 112
Db 47 RGGGGGGGDDGEGGCYLRRL-----PFVAAAVVSSRREALGDS 90
QY 113 LSRCSSTASSVNLGQGRGSHTCRSYDAA-----EAGGDHVLVDVSAASN 158
Db 91 VAEAAASSSSRAVELLGCSEEEAAEKVIDEAAWDVVVLQVCTQAGEDH--DEESSVGD 148
QY 159 SGSGPDRERETTPSSR---AHGELSLESD-----LAGHKTGSPSLPAAT 200
Db 149 SGCG--RERSATTSSRRPPGADSSDAESNQEAKQCMCRSSSTTSAAPHAG---ATT 202
QY 201 PAELIVPP--AHEIOEFFAAAEAAQAKRFASKYNFDFVRGVPLDAG--GRFEW 250
Db 203 RSFRMAPPAAAAEIEEFLAAAEERSEAEERFAAKYNFDFVRGVPLDAGGAGRFEW 256

RESULT 13

US-10-333-006-55
; Sequence 55, Application US/10333006
; Publication No. US20040019926A1
; GENERAL INFORMATION:

; APPLICANT: Frankard, Valerie Marie-No. US20040019926A111le S.
; APPLICANT: Peres Bota, Adrian Marius
; APPLICANT: Droual, Anne-Marie
; APPLICANT: Mironov, Vladimir
; APPLICANT: Inz, Dirk
; APPLICANT: Hatzfeld, Yves

; TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
; FILE REFERENCE: 1187-13

; CURRENT APPLICATION NUMBER: US/10/333,006
; CURRENT FILING DATE: 2003-01-14

; PRIOR APPLICATION NUMBER: PCT/IB01/01492
; PRIOR FILING DATE: 2001-06-29

; PRIOR APPLICATION NUMBER: US 60/218,471
; PRIOR FILING DATE: 2000-07-14

; PRIOR APPLICATION NUMBER: US 60/241,219
; PRIOR FILING DATE: 2000-10-13

; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 55
; LENGTH: 226

; TYPE: PRT

ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: probe or primer
FEATURE:
NAME/KEY: misc feature
LOCATION: (119)..(119)
OTHER INFORMATION: Xaa = any amino acid
US-10-333-006-55

Query Match 21.4%; Score 279; DB 15; Length 226;
Best Local Similarity 34.4%; Pred. No. 1.8e-15;
Matches 98; Conservative 26; Mismatches 71; Indels 90; Gaps 13;

QY 1 MGKYMRCRGAAGAEEVAEEVTQVVGVRTRSRSAATGGVAKVAPRRKRAPAGEPAAAVS 60
Db 1 MGK-KKRDGAAARRQARVVVG---GVRTR-----AAVTARRVVASAEEGCGLVG 46
QY 61 AGDGGG-----CYIHLRSRMLFMAPPQPQPSVDSVPTPVEAADGAAGQGAALAAAG 112
Db 47 RGGGGGGDDGEGGCYLRRLRRL-----PFVAAVVSSRREALGDS 90
QY 113 LSRCSSTASSVNLGLGQSGSHTCRSYDAAEAGGDHVLVDVSAASNSGSGPDRERETTP 172
Db 91 VAEAASSSSRAVELLG-----C-----SGEEHMAEKXSA-----TTP 124
QY 173 SSR---AHGELSDLESD-----LAGHKTGPSLPAATPAAELIVPP--AHE 212
Db 125 SSRPPGDADSSDAESNQEAQKQCMRRSSSTTSAAPFAG----ATTRSRMAPPAAAAE 180
QY 213 IQEFAAAEAQAQKRFASKYNFDFVRGVPLDAG--GRFEWAPVVS 255
Db 181 IEEFLAAERSEAEERFAAKYNFDFVRGVPLDAGGAGRFETAVGS 225

RESULT 14

US-10-437-963-120870
Sequence 120870, Application US/10437963
Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: Ia Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 120870

LENGTH: 194

TYPE: PRT

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT4530_23950C.1.pap

US-10-437-963-120870

Query Match 18.3%; Score 238; DB 16; Length 194;
Best Local Similarity 31.5%; Pred. No. 4.3e-12;
Matches 82; Conservative 22; Mismatches 76; Indels 80; Gaps 11;

QY 1 MGKYMRCRGAAGAEEVAEEVTQV-----VGVTRSRSAATGGVAKVAPRRKRAPAGEPA 56
Db 1 MGKYMRCRGAAGAEEVAEEVTQV-----VGVTRSRSAATGGVAKVAPRRKRAPAGEPA 56
QY 57 AAVSAGDGGSCYIHLRSRMLFMAPPQPQPSVDSVPTPVEAADGAAGQGAALAAAGL SRC 116
Db 44 -----QGGE-----YLELRSRRLKLELPPPPP-----PPRRRRATAAATAATAA----- 83

QY 117 SSTASSVNLGLGQSGSHTCRSYDAAEA-----GGDHVLVDVSAASNSGSGPDRERETTP 172
Db 84 -----ESAEAEVSFGGENVL-ELEAM-----ERNTRETTP 112
QY 173 SS--RAHGELSDLESDLAGHKTGPSLPAATPAAELIVPPAHEIQEFAAAEAQAARFAS 230
Db 113 CSLIRDPDTISTPGSTTRRSHSSHCKVQTPVRHNIIPASAELEAFAAEERQORQAFID 172
QY 231 KYNFDFVRGVPLDAGGRFEW 250
Db 173 KYNFDPVNDCLP--GRFEW 190

RESULT 15

US-10-333-006-44

Sequence 44, Application US/10333006

Publication No. US20040019926A1

GENERAL INFORMATION:

APPLICANT: Frankard, Valerie Marie-No. US20040019926A1lle S.
APPLICANT: Peres Bota, Adrian Marius
APPLICANT: Droual, Anne-Marie
APPLICANT: Mironov, Vladimír
APPLICANT: Inz, Dirk
APPLICANT: Hatzfeld, Yves

TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS

FILE REFERENCE: 1187-13

CURRENT APPLICATION NUMBER: US/10/333,006

CURRENT FILING DATE: 2003-01-14

PRIOR APPLICATION NUMBER: PCT/IB01/01492

PRIOR FILING DATE: 2001-06-29

PRIOR APPLICATION NUMBER: US 60/218,471

PRIOR FILING DATE: 2000-07-14

PRIOR APPLICATION NUMBER: US 60/241,219

NUMBER OF SEQ ID NOS: 59

SOFTWARE: PatentIn version 3.1

SEQ ID NO 44

LENGTH: 194

TYPE: PRT

ORGANISM: Oryza sativa

US-10-333-006-44

Query Match 18.0%; Score 235; DB 15; Length 194;
Best Local Similarity 31.5%; Pred. No. 7.7e-12;
Matches 82; Conservative 25; Mismatches 73; Indels 80; Gaps 12;

QY 1 MGKYMRCRGAAGAEEVAEEVTQV-----VGVTRSRSAATGGVAKVAPRRKRAPAGEPA 56

Db 1 MGKYMRCRGAAGAEEVAEEVTQV-----VGVTRSRSAATGGVAKVAPRRKRAPAGEPA 56

QY 57 AAVSAGDGGSCYIHLRSRMLFMAPPQPQPSVDSVPTPVEAADGAAGQGAALAAAGL SRC 116

Db 44 -----QGGE-----YLELRSRRLKLELPPPPP-----PPRRRRATAA-----ATA 78

QY 117 SSTASSVNLGLGQSGSHTCRSYDAAEA-----GGDHVLVDVSAASNSGSGPDRERETTP 172

Db 79 DATAT-----ESAEAEVSFGGENVL-ELEAM-----ERNTRETTP 112

QY 173 SS--RAHGELSDLESDLAGHKTGPSLPAATPAAELIVPPAHEIQEFAAAEAQAARFAS 230

Db 113 CSLIRDPDTISTPGSTTRRSHSSHCKVQTPVRHNIIPASAELEAFAAEERQORQAFID 172

QY 231 KYNFDFVRGVPLDAGGRFEW 250

Db 173 KYNFDPVNDCLP--GRFEW 190

Search completed: October 1, 2004, 17:05:50
Job time : 131 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 17:03:42 ; Search time 41 Seconds
(without alignments)
600.611 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 256
Sequence: 1 MGKYMVKCRGAAGAEVAAVE.....VRGVPLDAGGRFEWAPVVS I 256

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_78:*
1: Pir1:*
2: Pir2:*
3: Pir3:*
4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|----------|--------------------|
| 1 | 10 | 3.9 | 277 | 2 H84314 | Cytochrome aa3 con |
| 2 | 8 | 3.1 | 173 | 2 T45993 | hypothetical prote |
| 3 | 8 | 3.1 | 181 | 2 T45990 | hypothetical prote |
| 4 | 8 | 3.1 | 215 | 2 T03680 | plasma membrane pr |
| 5 | 8 | 3.1 | 272 | 1 S41498 | DNA excision repai |
| 6 | 8 | 3.1 | 334 | 2 T29061 | probable integral |
| 7 | 8 | 3.1 | 404 | 2 D70977 | hypothetical prote |
| 8 | 8 | 3.1 | 427 | 2 AC1627 | N-acetylmuramoyl-L |
| 9 | 8 | 3.1 | 427 | 2 AI1264 | N-acetylmuramoyl-L |
| 10 | 8 | 3.1 | 452 | 2 S03829 | nifB protein - Rho |
| 11 | 8 | 3.1 | 599 | 2 S65180 | hypothetical prote |
| 12 | 8 | 3.1 | 601 | 2 S12004 | tyramine receptor |
| 13 | 8 | 3.1 | 601 | 2 JH0170 | octopamine recepto |
| 14 | 8 | 3.1 | 1306 | 2 A70934 | hypothetical glyci |
| 15 | 7 | 2.7 | 44 | 1 BXS41 | antibacterial prot |
| 16 | 7 | 2.7 | 73 | 2 E70833 | hypothetical prote |
| 17 | 7 | 2.7 | 97 | 2 T18130 | hypothetical prote |
| 18 | 7 | 2.7 | 99 | 2 S59145 | NADH2 dehydrogenas |
| 19 | 7 | 2.7 | 101 | 2 T26641 | hypothetical prote |
| 20 | 7 | 2.7 | 118 | 2 S47024 | ribosomal protein |
| 21 | 7 | 2.7 | 118 | 2 AF1260 | hypothetical prote |
| 22 | 7 | 2.7 | 118 | 2 AH1622 | hypothetical prote |
| 23 | 7 | 2.7 | 125 | 2 H70659 | hypothetical prote |
| 24 | 7 | 2.7 | 125 | 2 F72785 | hypothetical prote |
| 25 | 7 | 2.7 | 126 | 2 F84299 | hypothetical prote |
| 26 | 7 | 2.7 | 130 | 2 A83400 | hypothetical prote |
| 27 | 7 | 2.7 | 131 | 2 B84598 | hypothetical prote |
| 28 | 7 | 2.7 | 132 | 2 T35141 | hypothetical prote |
| 29 | 7 | 2.7 | 134 | 1 PWQFE | H+-transporting tw |

| | | | | | |
|----|---|-----|-----|----------|--------------------|
| 30 | 7 | 2.7 | 139 | 2 A26892 | Mopa box protein - |
| 31 | 7 | 2.7 | 140 | 2 S27658 | hypothetical prote |
| 32 | 7 | 2.7 | 142 | 2 T46542 | hypothetical prote |
| 33 | 7 | 2.7 | 142 | 2 S50662 | hypothetical prote |
| 34 | 7 | 2.7 | 152 | 1 MYPN | myoglobin - empero |
| 35 | 7 | 2.7 | 153 | 2 JT0636 | myoglobin - rhinoc |
| 36 | 7 | 2.7 | 153 | 2 JC5190 | myoglobin - rhinoc |
| 37 | 7 | 2.7 | 153 | 2 JC7789 | myoglobin - Japane |
| 38 | 7 | 2.7 | 153 | 2 JC7791 | myoglobin - thick- |
| 39 | 7 | 2.7 | 160 | 2 B70718 | hypothetical prote |
| 40 | 7 | 2.7 | 163 | 2 E75416 | hypothetical prote |
| 41 | 7 | 2.7 | 163 | 2 F84651 | probable kinetecho |
| 42 | 7 | 2.7 | 163 | 2 B86335 | hypothetical prote |
| 43 | 7 | 2.7 | 185 | 2 B70072 | conserved hypothe |
| 44 | 7 | 2.7 | 187 | 2 H81347 | probable decarboxy |
| 45 | 7 | 2.7 | 192 | 2 T26386 | hypothetical prote |

ALIGNMENTS

RESULT 1

H84314
cytochrome aa3 controlling protein [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: H84314
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.;
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Lie
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: H84314
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-277 <STO>
A;Cross-references: GB:AE004437; NID:g10581096; PIDN:AAG19884.1; GSPDB:GN00138
C;Genetics:
A;Gene: ccp

Query Match 3.9%; Score 10; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 R5AAATGGVA 41
|||||
Db 216 R5AAATGGVA 225

RESULT 2

T45993
hypothetical protein F9D24.240 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C;Accession: T45993
R;D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.;
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23011
A;Accession: T45993
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-173 <DAN>
A;Cross-references: EMBL:AL137081
A;Experimental source: cultivar Columbia; BAC clone F9D24
C;Genetics:
A;Map position: 3
A;Introns: 41/3
A;Note: F9D24.240

Query Match 3.1%; Score 8; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 12;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 LSDLES DL 187
| | | | | | | |
Db 148 LSDLES DL 155

RESULT 3
T45990
hypothetical protein F9D24.210 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 28-Jul-2000
C;Accession: T45990
R;D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.;
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23011
A;Accession: T45990
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-181 <DAN>
A;Cross-references: EMBL:AL137081
A;Experimental source: cultivar Columbia; BAC clone F9D24
C;Genetics:
A;Map position: 3
A;Introns: 47/3
A;Note: F9D24.210
C;Superfamily: Arabidopsis thaliana hypothetical protein F9D24.210

Query Match 3.1%; Score 8; DB 2; Length 181;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 LSDLES DL 187
| | | | | | | |
Db 154 LSDLES DL 161

RESULT 4
T03680
plasma membrane protein - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 20-Jun-2000
C;Accession: T03680
R;Logan, D.C.; Domergue, O.; Teyssendier de la Serve, B.; Rossignol, M.
submitted to the EMBL Data Library, October 1996
A;Description: A new family of plasma membrane intrinsic polypeptides differentially reg
A;Reference number: Z15001
A;Accession: T03680
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-215 <LOG>
A;Cross-references: EMBL:Y08609
A;Experimental source: strain xanthi N.C; tissue-type leaf

Query Match 3.1%; Score 8; DB 2; Length 215;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 PAATPAAE 204
| | | | | | | |
Db 176 PAATPAAE 183

RESULT 5
S41498
DNA excision repair protein XPAC - mouse
N;Alternate names: Xeroderma pigmentosum group A-complementing protein
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S41498
R;van Oostrom, C.T.M.; de Vries, A.; Verbeek, S.J.; van Kreijl, C.F.; van Steeg, H.
Nucleic Acids Res. 22, 11-14, 1994
A;Title: Cloning and characterization of the mouse XPAC gene.

A;Reference number: S41498; MUID:94173654; PMID:8127648
A;Accession: S41498
A;Status: preliminary; translation not shown
A;Molecule type: mRNA
A;Residues: 1-272 <VAN>
A;Cross-references: EMBL:X74351; NID:g440564; PIDN:CAA52393.1; PID:g440565
C;Superfamily: DNA excision repair protein XPAC
C;Keywords: DNA binding; zinc finger

Query Match 3.1%; Score 8; DB 1; Length 272;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AAATGGVA 41
| | | | | | | |
Db 53 AAATGGVA 60

RESULT 6
T29061
probable integral membrane protein - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C;Accession: T29061
R;Redenbach, M.; Kieser, H.M.; Denapaita, D.; Eichner, A.; Cullum, J.; Kinashi, H.; Hopw
Mol. Microbiol. 21, 77-96, 1996
A;Title: A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb St
A;Reference number: Z20556; MUID:97000351; PMID:8843436
A;Accession: T29061
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-334 <RED>
A;Cross-references: EMBL:AL031124; NID:el312893; PID:el312905; PIDN:CAA19979.1
C;Genetics:
A;Note: SCIC2.12c

Query Match 3.1%; Score 8; DB 2; Length 334;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 GAAGAEVA 17
| | | | | | | |
Db 141 GAAGAEVA 148

RESULT 7
D70977
hypothetical protein Rv3446c - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: D70977
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70977
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-404 <COL>
A;Cross-references: GB:Z95389; GB:AL123456; NID:g3242261; PIDN:CAB08676.1; PID:e315967;
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: Rv3446c

Query Match 3.1%; Score 8; DB 2; Length 404;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 AGAEVAAV 19
| | | | | | | |

Db 131 AGAEVAAV 138

RESULT 8

AC1627

N-acetylmuramoyl-L-alanine amidase homolog lin1556 [imported] - Listeria innocua (strain C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AC1627
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AC1627
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-427 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC96787.1; PID:gl16414043; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin1556

Query Match

Best Local Similarity 3.1%; Score 8; DB 2; Length 427;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 DVSAASNS 159

|||||

Db 92 DVSAASNS 99

RESULT 9

AI1264

N-acetylmuramoyl-L-alanine amidase homolog lmo1521 [imported] - Listeria monocytogenes C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AI1264
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AI1264
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-427 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAC99599.1; PID:gl16410950; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo1521

Query Match

Best Local Similarity 3.1%; Score 8; DB 2; Length 427;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 DVSAASNS 159

|||||

Db 92 DVSAASNS 99

RESULT 10

S03829

nifB protein - Rhodobacter capsulatus

C;Species: Rhodobacter capsulatus

C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 21-Jul-2000

C;Accession: S03829; A36914

R;Masepohl, B.; Klipp, W.; Puehler, A.

Mol. Gen. Genet. 212, 27-37, 1988

A;Title: Genetic characterization and sequence analysis of the duplicated nifA/nifB gene

A;Reference number: S03828; MUID:88232432; PMID:2836706

A;Accession: S03829

A;Molecule type: DNA

A;Residues: 1-452 <MAS>

A;Cross-references: EMBL:X07567; NID:g5748652; PIDN:CAA30450.1; PID:g46063

R;Wang, G.; Angermuller, S.; Klipp, W.

J. Bacteriol. 175, 3031-3042, 1993

A;Title: Characterization of Rhodobacter capsulatus genes encoding a molybdenum transport

A;Reference number: A36914; MUID:93259949; PMID:8491722

A;Accession: A36914

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 381-452 <WAN>

A;Note: sequence extracted from NCBI backbone (NCBIN:131915, NCBIP:131916)

C;Genetics:

A;Gene: nifB

C;Superfamily: Rhizobium nifB protein

Query Match

Best Local Similarity 3.1%; Score 8; DB 2; Length 452;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 AAAEAAQA 225

|||||

Db 316 AAAEAAQA 323

RESULT 11

S65180

hypothetical protein YPL169c - Yeast (Saccharomyces cerevisiae)

N;Alternate names: hypothetical protein P2520

C;Species: Saccharomyces cerevisiae

C;Date: 10-Dec-1994 #sequence_revision 31-May-1996 #text_change 19-Apr-2002

C;Accession: S65180; S69430

R;Purnelle, B.; Coster, F.; Goffeau, A.

submitted to the Protein Sequence Database, May 1996

A;Reference number: S65154

A;Accession: S65180

A;Molecule type: DNA

A;Residues: 1-599 <PUR>

A;Cross-references: EMBL:Z73525; NID:gl370356; PID:gl370357; MIPS:YPL169c

A;Experimental source: strain S288C (AB972)

R;Purnelle, B.; Comblez, S.; Coster, F.; Naveau, F.; Goffeau, A.

submitted to the EMBL Data Library, March 1996

A;Description: The sequence of 55 kb on the left arm of yeast chromosome XVI identifies

ogues to the human phosphotyrosyl phosphatase activator PTPA and a homologue to the plant

A;Reference number: S69428

A;Accession: S69430

A;Molecule type: DNA

A;Residues: 1-599 <PUW>

A;Cross-references: EMBL:X96770; NID:gl403537; PID:gl403540

C;Genetics:

A;Gene: SGD:MEX67

A;Cross-references: SGD:S0006090

A;Map position: 16L

Query Match

Best Local Similarity 3.1%; Score 8; DB 2; Length 599;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 PPQPQPSV 88

|||||

Db 506 PPQPQPSV 513

RESULT 12

S12004

tyramine receptor - fruit fly (Drosophila sp.)

C;Species: Drosophila sp.

C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 04-Sep-1998

C;Accession: S12004
R;Saudou, F.; Amlaiky, N.; Plassat, J.L.; Borrelli, E.; Hen, R.
EMBO J. 9, 3611-3617, 1990
A;Title: Cloning and characterization of a Drosophila tyramine receptor.
A;Reference number: S12004; MUID:91006061; PMID:2170118
A;Accession: S12004
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-601 <SAU>
C;Superfamily: octopamine receptor type I

Query Match 3.1%; Score 8; DB 2; Length 601;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 AGDGGGSC 68
| | | | | | | |
Db 435 AGDGGGSC 442

RESULT 13
JH0170
octopamine receptor type I - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 10-Dec-1999
C;Accession: JH0170
R;Arakawa, S.; Gocayne, J.D.; McCombie, W.R.; Urquhart, D.A.; Hall, L.M.; Fraser, C.M.;
Neuron 4, 343-354, 1990
A;Title: Cloning, localization, and permanent expression of a Drosophila octopamine rece
A;Reference number: JH0170; MUID:90198291; PMID:2156539
A;Accession: JH0170
A;Molecule type: mRNA
A;Residues: 1-601 <ARA>
A;Cross-references: GB:M60789; GB:M26181; NID:g158002; PIDN:AAA28731.1; PID:g158003
A;Experimental source: strain oregon-R
A;Note: this sequence shows homology with alpha-2 adrenergic receptors
C;Genetics:
A;Gene: FlyBase:Ocr
A;Cross-references: FlyBase:FBgn0004514
C;Superfamily: octopamine receptor type I
C;Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F;110-133/Domain: transmembrane #status predicted <TM1>
F;147-169/Domain: transmembrane #status predicted <TM2>
F;184-207/Domain: transmembrane #status predicted <TM3>
F;227-250/Domain: transmembrane #status predicted <TM4>
F;267-290/Domain: transmembrane #status predicted <TM5>
F;529-552/Domain: transmembrane #status predicted <TM6>
F;562-585/Domain: transmembrane #status predicted <TM7>
F;11,57,324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 3.1%; Score 8; DB 2; Length 601;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 AGDGGGSC 68
| | | | | | | |
Db 435 AGDGGGSC 442

RESULT 14
A70934
hypothetical glycine-rich protein Rv0578c - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C;Accession: A70934
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqaes, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: A70934

A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1306 <COL>
A;Cross-references: GB:AL021942; GB:AL123456; NID:g3242298; PIDN:CAA17449.1; PID:g290963;
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: Rv0578c
C;Superfamily: collagen alpha 1(IV) chain

Query Match 3.1%; Score 8; DB 2; Length 1306;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 VSAGGDGG 66
| | | | | | | |
Db 1191 VSAGGDGG 1198

RESULT 15
BXSA1
antibacterial protein 1 - Staphylococcus haemolyticus
N;Alternate names: gonococcal growth inhibitor 1
C;Species: Staphylococcus haemolyticus
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 24-Nov-1999
C;Accession: S00599
R;Watson, D.C.; Yaguchi, M.; Bisailon, J.G.; Beaudet, R.; Morosoli, R.
Biochem. J. 252, 87-93, 1988
A;Title: The amino acid sequence of a gonococcal growth inhibitor from Staphylococcus ha
A;Reference number: S00599; MUID:88339821; PMID:3138972
A;Accession: S00599
A;Molecule type: protein
A;Residues: 1-44 <WAT>
C;Superfamily: Staphylococcus haemolyticus antibacterial protein
C;Keywords: antibacterial; antibiotic; blocked amino end; hemolysis
F;1/Modified site: blocked amino end (Met) (probably formylated) #status experimental

Query Match 2.7%; Score 7; DB 1; Length 44;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 AAASVSAG 62
| | | | | | | |
Db 9 AAASVSAG 15

Search completed: October 1, 2004, 17:11:20
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 16:55:17 ; Search time 24 Seconds
(without alignments)
555.415 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 256
Sequence: 1 MGKYMCKRGAGAEVAAVE.....VRGVPLDAGGRFEWAPVVS 256

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|--------------|---------------------|
| 1 | 8 | 3.1 | 250 | 1 LINK_PSEPA | P50198 pseudomonas |
| 2 | 8 | 3.1 | 272 | 1 XPA_MOUSE | Q64267 mus musculus |
| 3 | 8 | 3.1 | 352 | 1 DMPF_ACICA | Q7wtj2 acinetobact |
| 4 | 8 | 3.1 | 452 | 1 NIFB_RHOCA | P17434 rhodobacter |
| 5 | 8 | 3.1 | 492 | 1 TMS2_HUMAN | O15393 homo sapien |
| 6 | 8 | 3.1 | 599 | 1 MX67_YEAST | Q99257 saccharomyc |
| 7 | 8 | 3.1 | 601 | 1 OAR_DROME | P22270 drosophila |
| 8 | 7 | 2.7 | 44 | 1 GGI1_STAHA | P11697 staphylococ |
| 9 | 7 | 2.7 | 99 | 1 NULM_ALBCO | P48928 albinaria c |
| 10 | 7 | 2.7 | 114 | 1 H33_STYLE | P81197 stylonychia |
| 11 | 7 | 2.7 | 114 | 1 H36_STYLE | P81200 stylonychia |
| 12 | 7 | 2.7 | 114 | 1 H37_STYLE | P81201 stylonychia |
| 13 | 7 | 2.7 | 114 | 1 H39_STYLE | P81196 stylonychia |
| 14 | 7 | 2.7 | 118 | 1 R18E_SULAC | P39474 sulfolobus |
| 15 | 7 | 2.7 | 124 | 1 CRCB_PSEPK | Q88ft1 pseudomonas |
| 16 | 7 | 2.7 | 134 | 1 ATPE_RHORU | P05442 rhodospiril |
| 17 | 7 | 2.7 | 142 | 1 NCBI_YEAST | P40096 saccharomyc |
| 18 | 7 | 2.7 | 149 | 1 YCIA_ZYMMO | O66120 zymomonas m |
| 19 | 7 | 2.7 | 152 | 1 MYG_APTFO | P02199 aptenodytes |
| 20 | 7 | 2.7 | 160 | 1 Y964_MYCTU | P71546 mycobacteri |
| 21 | 7 | 2.7 | 160 | 1 Y989_MYCBO | P59978 mycobacteri |
| 22 | 7 | 2.7 | 187 | 1 PAAD_CAMJE | Q9ppf1 campylobact |
| 23 | 7 | 2.7 | 208 | 1 COBH_MYCTU | Q10676 mycobacteri |
| 24 | 7 | 2.7 | 220 | 1 Y085_MYCTU | Q10882 mycobacteri |
| 25 | 7 | 2.7 | 230 | 1 CUT1_FUSSO | P00590 fusarium so |
| 26 | 7 | 2.7 | 230 | 1 CUTI_FUSSC | Q99174 fusarium so |
| 27 | 7 | 2.7 | 231 | 1 CUT3_FUSSO | Q96us9 fusarium so |
| 28 | 7 | 2.7 | 242 | 1 YK81_STRCO | P45497 streptomyc |
| 29 | 7 | 2.7 | 263 | 1 GNP2_GIALA | O97440 giardia lam |
| 30 | 7 | 2.7 | 289 | 1 AROK_METH | O26896 methanobact |
| 31 | 7 | 2.7 | 296 | 1 Y4YQ_RHISN | P55725 rhizobium s |
| 32 | 7 | 2.7 | 327 | 1 THIO_RHIET | O34292 rhizobium e |
| 33 | 7 | 2.7 | 335 | 1 CPR4_CAEEL | P43508 caenorhabdi |

| | | | | | |
|----|---|-----|-----|--------------|--------------------|
| 34 | 7 | 2.7 | 337 | 1 COBU_PSEDE | P29935 pseudomonas |
| 35 | 7 | 2.7 | 338 | 1 COBT_RHIME | Q92p98 rhizobium m |
| 36 | 7 | 2.7 | 342 | 1 COBT_METSQ | Q9x7f4 methylobact |
| 37 | 7 | 2.7 | 344 | 1 COBT_AGR5 | Q8ue71 agrobacteri |
| 38 | 7 | 2.7 | 349 | 1 MOAA_RHIME | Q92pb4 rhizobium m |
| 39 | 7 | 2.7 | 353 | 1 AOX1_TOBAC | Q41224 nicotiana t |
| 40 | 7 | 2.7 | 361 | 1 IF35_MOUSE | Q9dch4 mus musculu |
| 41 | 7 | 2.7 | 371 | 1 PGLR_PENJA | O42824 penicillium |
| 42 | 7 | 2.7 | 385 | 1 VE2_COPV | Q89420 canine oral |
| 43 | 7 | 2.7 | 394 | 1 ATOB_ECOLI | P76461 escherichia |
| 44 | 7 | 2.7 | 398 | 1 THIL_YEAST | P41338 saccharomyc |
| 45 | 7 | 2.7 | 406 | 1 FBW3_HUMAN | Q9ukb7 homo sapien |

ALIGNMENTS

RESULT 1

| ID | LINK_PSEPA | STANDARD | PRT | 250 AA |
|----|--|----------|-----|--------|
| AC | P50198; | | | |
| DT | 01-OCT-1996 (Rel. 34, Created) | | | |
| DT | 01-OCT-1996 (Rel. 34, Last sequence update) | | | |
| DT | 10-OCT-2003 (Rel. 42, Last annotation update) | | | |
| DE | 2,5-dichloro-2,5-cyclohexadiene-1,4-diol dehydrogenase (EC 1.1.1.1.-) | | | |
| DE | (2,5-DDOL dehydrogenase). | | | |
| GN | LINK. | | | |
| OS | Pseudomonas paucimobilis (Sphingomonas paucimobilis). | | | |
| OC | Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales; | | | |
| OC | Sphingomonadaceae; Sphingomonas. | | | |
| OX | NCBI_TaxID=13689; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=UT26; | | | |
| RX | MEDLINE=94252977; PubMed=7515041; | | | |
| RA | Nagata Y., Ohtomo R., Miyauchi K., Fukuda M., Yano K., Takagi M.; | | | |
| RT | "Cloning and sequencing of a 2,5-dichloro-2,5-cyclohexadiene-1,4-diol | | | |
| RT | dehydrogenase gene involved in the degradation of gamma- | | | |
| RT | hexachlorocyclohexane in Pseudomonas paucimobilis."; | | | |
| RL | J. Bacteriol. 176:3117-3125(1994). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=B90; | | | |
| RX | MEDLINE=22338262; PubMed=12450824; | | | |
| RA | Kumari R., Subudhi S., Suar M., Dhinagra G., Raina V., Dogra C., | | | |
| RT | Lal S., van der Meer J.R., Holliger C., Lal R.; | | | |
| RT | "Cloning and characterization of lin genes responsible for the | | | |
| RT | degradation of hexachlorocyclohexane isomers by Sphingomonas | | | |
| RT | paucimobilis strain B90."; | | | |
| RL | Appl. Environ. Microbiol. 68:6021-6028(2002). | | | |
| CC | -!- FUNCTION: Catalyzes the degradation of 2,5-dichloro-2,5- | | | |
| CC | cyclohexadiene-1,4-diol (2,5-DDOL) into 2,5-dichlorohydroquinone | | | |
| CC | (2,5-DCHQ). Link appears not to be involved in gamma-HCH | | | |
| CC | degradation pathway. | | | |
| CC | -!- CATALYTIC ACTIVITY: 2,5-DDOL + NAD(+) = 2,5-DCHQ + NADH. | | | |
| CC | -!- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases | | | |
| CC | (SDR) family. | | | |
| CC | ----- | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | | |
| CC | or send an email to license@isb-sib.ch). | | | |
| CC | ----- | | | |
| DR | EMBL; D23722; BAA04939.1; -. | | | |
| DR | EMBL; AY150579; AAN64237.1; -. | | | |
| DR | HSSP; P19992; 1HDC. | | | |
| DR | InterPro; IPR002198; ADH_short. | | | |
| DR | Pfam; PF00106; adh_short; 1. | | | |
| DR | PRINTS; PR00080; SDRFAMILY. | | | |
| DR | PROSITE; PS00061; ADH_SHORT; 1. | | | |

KW Aromatic hydrocarbons catabolism; Oxidoreductase; NAD.
FT NP BIND 9 34 NAD (BY SIMILARITY).
FT ACT_SITE 156 156 BY SIMILARITY.
FT VARIANT 80 81 DG -> EA (IN STRAIN B90).
SQ SEQUENCE 250 AA; 25492 MW; 8C52703FF76382CF CRC64;
Query Match 3.1%; Score 8; DB 1; Length 250;
Best Local Similarity 100.0%; Pred.No. 7.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 222 AAQAKRFA 229
Db 21 AAQAKRFA 28
RESULT 2
XPA_MOUSE STANDARD; PRT; 272 AA.
AC Q64267;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA-repair protein complementing XP-A cells homolog (Xeroderma pigmentosum group A complementing protein homolog).
GN XPA OR XPAC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Ola;
RX MEDLINE=94173654; PubMed=8127648;
RA van Oostrom C.T.M., de Vries A., Verbeek S.J., van Kreijl C.F.,
RA van Steeg H.;
RT "Cloning and characterization of the mouse XPAC gene."
RL Nucleic Acids Res. 22:11-14(1994).
CC -!- FUNCTION: Involved in DNA excision repair. Initiates repair by
CC binding to damaged sites with various affinities, depending on the
CC photoproduct and the transcriptional state of the region (By
CC similarity).
CC -!- SUBUNIT: Interacts with XAB1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the XPA family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X74345; CAA52392.1; --
CC EMBL; X74346; CAA52392.1; JOINED.
CC EMBL; X74347; CAA52392.1; JOINED.
CC EMBL; X74348; CAA52392.1; JOINED.
CC EMBL; X74349; CAA52392.1; JOINED.
CC EMBL; X74350; CAA52392.1; JOINED.
CC EMBL; X74351; CAA52393.1; --
CC PIR; S41498; S41498.
CC MGD; MGI:99135; Xpa.
CC GO; GO:0006289; P:nucleotide-excision repair; IMP.
CC GO; GO:0006979; P:response to oxidative stress; IMP.
CC InterPro; IPR000465; XPA_protein.
CC Pfam; PF01286; XPA_N; 1.
CC TIGRfams; TIGR00598; rad14; 1.
CC PROSITE; PS00752; XPA_1; 1.
CC PROSITE; PS00753; XPA_2; 1.
KW DNA repair; DNA-binding; Zinc-finger; Nuclear protein.
FT DOMAIN 28 49 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 78 82 POLY-GLU.

FT ZN FING 104 128
SQ SEQUENCE 272 AA; 31399 MW; 9061C952817E8DCB CRC64;
Query Match 3.1%; Score 8; DB 1; Length 272;
Best Local Similarity 100.0%; Pred.No. 7.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 34 AAATGGVA 41
Db 53 AAATGGVA 60
RESULT 3
DMPP_ACICA STANDARD; PRT; 352 AA.
ID DMPP_ACICA
AC Q7WTJ2;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phenol hydroxylase P5 protein (EC 1.14.13.7) (Phenol 2-monooxygenase
DE P5 component).
GN MPH.
OS Acinetobacter calcoaceticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=471;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PHEA-2;
RX MEDLINE=22617848; PubMed=12732969;
RA Xu Y., Chen M., Zhang W., Lin M.;
RT "Genetic organization of genes encoding phenol hydroxylase, benzoate
RT 1,2-dioxygenase alpha subunit and its regulatory proteins in
RT Acinetobacter calcoaceticus PHEA-2."
RL Curr. Microbiol. 46:235-240(2003).
RN [2]
RP SEQUENCE OF 1-25, AND INDUCTION.
RC STRAIN=69-V;
RA Benndorf D., Davidson I., Babel W.;
RT "Growth on phenol at chemotress levels enhances induction of the
RT phenol degradation pathway in Acinetobacter calcoaceticus.";
RL Microbiology 0:0-0(2003).
CC -!- FUNCTION: Catabolizes phenol, and some of its methylated
CC derivatives. P5 is required for growth on phenol, and for in vitro
CC phenol hydroxylase activity (By similarity).
CC -!- FUNCTION: Probable electron transfer from NADPH, via FAD and the
CC 2Fe-2S center, to the oxygenase activity site of the enzyme (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: Phenol + NADPH + O(2) = catechol + NADP(+) +
CC H(2)O.
CC -!- COFACTOR: P5 contains a FAD cofactor, and a ferredoxin-type
CC [2Fe-2S] center (By similarity).
CC -!- PATHWAY: Phenol biodegradation; first step.
CC -!- SUBUNIT: The multicomponent enzyme phenol hydroxylase is formed by
CC P0, P1, P2, P3, P4 and P5 polypeptides (By similarity).
CC -!- INDUCTION: By phenol.
CC -!- SIMILARITY: In the N-terminal section; belongs to the 2Fe2S plant-
CC type ferredoxin family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AJ564846; CAD92316.1; --
CC InterPro; IPR006058; 2Fe2S fd BS.
CC InterPro; IPR008333; FAD_binding_6.
CC InterPro; IPR001041; Ferredoxin.
CC InterPro; IPR010205; NqrF.
CC InterPro; IPR001433; Oxred_FAD/NAD(P).

DR Pfam; PF00970; FAD_binding_6; 1.
DR Pfam; PF00111; fer2; 1.
DR Pfam; PF00175; NAD_binding_1; 1.
DR TIGRFAMS; TIGR02008; fdx_p1ant; 1.
DR TIGRFAMS; TIGR01941; ngrF; 1.
DR PROSITE; PS00197; 2FE2S_FERREDOXIN; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase;
KW Monooxygenase; Flavoprotein; FAD; NADP;
KW Metal-binding; Iron-sulfur; Iron; 2Fe-2S;
KW Electron transport.
FT INIT_MET 0 0
FT METAL 36 36 IRON-SULFUR (2FE-2S) (BY
FT SIMILARITY).
FT METAL 41 41 IRON-SULFUR (2FE-2S) (BY
FT SIMILARITY).
FT METAL 44 44 IRON-SULFUR (2FE-2S) (BY
FT SIMILARITY).
FT METAL 76 76 IRON-SULFUR (2FE-2S) (BY
FT SIMILARITY).
SQ SEQUENCE 352 AA; 38906 MW; D15C8632F5620B7D CRC64;

Query Match 3.1%; Score 8; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 AADGAAGQ 104
|||||
Db 336 AADGAAGQ 343

RESULT 4
NIFB_RHOCA
ID_NIFB_RHOCA STANDARD; PRT; 452 AA.
AC P17434;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE FeMo cofactor biosynthesis protein nifB.
GN NIFB.
OS Rhodobacter capsulatus (Rhodospseudomonas capsulata).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Rhodobacter.
OX NCBI_TaxID=1061;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88232432; PubMed=2836706;
RA Masepohl P., Klipp W., Puehler A.;
RT "Genetic characterization and sequence analysis of the duplicated
nifA/nifB gene region of Rhodobacter capsulatus.";
RL Mol. Gen. Genet. 212:27-37(1988).
CC -!- FUNCTION: PROBABLY INVOLVED IN THE SYNTHESIS OF THE FE-MO
COFACTOR.
CC -!- SIMILARITY: BELONGS TO THE MOAA / NIFB / PQQE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X07567; CAA30450.1; -
DR PIR; S03829; S03829.
DR InterPro; IPR003731; DUF153.
DR InterPro; IPR006638; Elp3.
DR InterPro; IPR000385; Moaa_NifB_PqqE.
DR InterPro; IPR005980; NifB.
DR InterPro; IPR007197; Radical_SAM.
DR Pfam; PF02579; Nitro_FeMo-Co; 1.
DR Pfam; PF04055; Radical_SAM; 1.
DR SMART; SM00729; Elp3; 1.
DR TIGRFAMS; TIGR01290; nifB; 1.

DR PROSITE; PS01305; MOAA_NIFB_PQQE; 1.
KW Nitrogen fixation; Iron-sulfur.
FT METAL 34 34 IRON-SULFUR (POTENTIAL).
FT METAL 38 38 IRON-SULFUR (POTENTIAL).
FT METAL 41 41 IRON-SULFUR (POTENTIAL).
SQ SEQUENCE 452 AA; 49445 MW; EB6A7FEE31DDAD7D CRC64;

Query Match 3.1%; Score 8; DB 1; Length 452;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 AAAEAAQA 225
|||||
Db 316 AAAEAAQA 323

RESULT 5
TMS2_HUMAN
ID_TMS2_HUMAN STANDARD; PRT; 492 AA.
AC O15393; Q9BXX1;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transmembrane protease, serine 2 precursor (EC 3.4.21.-).
GN TMPRSS2 OR PRSS10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97468144; PubMed=9325052;
RA Paoloni-Giacobino A., Chen H., Peitsch M.C., Rossier C.,
RA Antonarakis S.E.;
RT "Cloning of the TMPRSS2 gene, which encodes a novel serine protease
with transmembrane, LDLRA, and SRCR domains and maps to 21q22.3.";
RL Genomics 44:309-320(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21309069; PubMed=11414763;
RA Teng D.H., Chen Y., Lian L., Ha P.C., Tavtigian S.V., Wong A.K.;
RT "Mutation analyses of 268 candidate genes in human tumor cell lines.";
RL Genomics 74:352-364(2001).
RN [3]
RP SEQUENCE FROM N.A., AND MUTAGENESIS.
RX MEDLINE=21139112; PubMed=11245484;
RA Afar D.E.H., Vivanco I., Hubert R.S., Kuo J., Chen E., Saffran D.C.,
RA Raitano A.B., Jakobovits A.;
RT "Catalytic cleavage of the androgen-regulated TMPRSS2 protease results
in its secretion by prostate and prostate cancer epithelia.";
RL Cancer Res. 61:1686-1692(2001).
RN [4]
RP TISSUE SPECIFICITY.
RX MEDLINE=21104370; PubMed=11169526;
RA Vaarala M.H., Porvari K.S., Kellokumpu S., Kyllonen A.P., Vihko P.T.;
RT "Expression of transmembrane serine protease TMPRSS2 in mouse and
human tissues.";
RL J. Pathol. 193:134-140(2001).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Activated by
cleavage and secreted.
CC -!- TISSUE SPECIFICITY: Expressed strongly in small intestine. Also
expressed in prostate, colon, stomach and salivary gland.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 1 LDL-receptor class A domain.
CC -!- SIMILARITY: Contains 1 SRCR domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).


```
CC -!- SIMILARITY: Belongs to the NXF family.
CC -!- SIMILARITY: Contains 4 leucine-rich (LRR) repeats.
CC -!- SIMILARITY: Contains 1 NTF2 domain.
CC -!- SIMILARITY: Contains 1 UBA-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X96770; CAA65552.1; -
CC EMBL; Z73525; CAA97875.1; -
CC PIR; S65180; S65180.
CC GerMOnline; 144151; -
CC SGD; S0006090; MEX67.
CC GO; GO:0005737; C:cytoplasm; IDA.
CC GO; GO:0005643; C:nuclear pore; IDA.
CC GO; GO:0005515; F:protein binding; IPI.
CC GO; GO:0003723; F:RNA binding; IDA.
CC GO; GO:0017056; F:structural constituent of nuclear pore; IDA.
CC GO; GO:0006406; P:mRNA-nucleus export; IMP.
CC InterPro; IPR002075; NTF2.
CC InterPro; IPR005637; TAP_C.
CC Pfam; PF03943; TAP_C; 1.
CC PROSITE; PS50177; NTF2 DOMAIN; 1.
KW Transport; mRNA transport; Nuclear protein; Repeat;
KW Leucine-rich repeat.
FT REPEAT 163 188 LRR 1.
FT REPEAT 189 214 LRR 2.
FT REPEAT 215 238 LRR 3.
FT REPEAT 239 264 LRR 4.
FT DOMAIN 280 467 NTF2.
FT DOMAIN 547 587 UBA-LIKE (BY SIMILARITY).
FT MUTAGEN 400 400 H->Y: IMPAIRS ASSOCIATION WITH THE
SQ SEQUENCE 599 AA; 67351 MW; 6CE55EEG6FA40B40 CRC64;

Query Match
Best Local Similarity 3.1%; Score 8; DB 1; Length 599;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 PPQPQPSV 88
Db 506 PPQPQPSV 513

RESULT 7
OAR_DROME STANDARD; PRT; 601 AA.
AC P22770; Q9VNW3;
DT 01-AUG-1991 (Rel. 19, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Tyramine/octopamine receptor precursor (Tyr/Oct-Dro).
GN TYRR OR OCR OR TYR OR OCTYR99AB OR CG7485.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Oregon-R;
RX MEDLINE=90198291; PubMed=2156539;
RA Arakawa S., Gocayne J.D., McCombie W.R., Urquhart D.A., Hall L.M.,
RT "Cloning, localization, and permanent expression of a Drosophila
RT octopamine receptor.";
RL Neuron 4:343-354(1990).
RN [2]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=Oregon-R; TISSUE=Head;
RX MEDLINE=91006061; PubMed=2170118;
RA Saudou F., Amlaiky N., Plassat J.-L., Borelli E., Hen R.;
RT "Cloning and characterization of a Drosophila tyramine receptor.";
RL EMBO J. 9:3611-3617(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -!- FUNCTION: Receptor for both octopamine and tyramine, invertebrate
CC neurotransmitters, and neuromodulators. The activity of this
CC receptor is mediated by G proteins which activate adenylyl
CC cyclase. The rank order of potency for agonists is tyramine >
CC octopamine > dopamine > epinephrine > norepinephrine > serotonin >
CC histamine. For antagonists, the rank order is yohimbine >
CC chlorpromazine > phentolamine > mianserine > cyproheptadine >
CC dihydroergotamine > clonidine > synephrine.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Preferentially expressed in Drosophila heads.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M60789; AAA28731.1; -
CC EMBL; X54794; CAA38565.1; -
CC EMBL; AE003596; AAF51802.1; -
CC PIR; S12004; S12004.
CC HSSP; P29274; 1MMH.
DR
```


DR FlyBase; FBgn0004514; TyrR.
DR GO; GO:0007608; P:olfaction; IMP.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCRHDOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 601 TYRAMINE/OCTOPAMINE RECEPTOR.
FT DOMAIN 27 112 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 113 135 1 (POTENTIAL).
FT DOMAIN 136 145 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 146 167 2 (POTENTIAL).
FT DOMAIN 168 184 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 185 205 3 (POTENTIAL).
FT DOMAIN 206 225 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 226 248 4 (POTENTIAL).
FT DOMAIN 249 273 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 274 295 5 (POTENTIAL).
FT DOMAIN 296 330 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 331 352 6 (POTENTIAL).
FT DOMAIN 353 366 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 367 388 7 (POTENTIAL).
FT DOMAIN 389 401 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 11 11 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 182 261 BY SIMILARITY.
FT CONFLICT 34 34 S -> N (IN REF. 1).
SQ SEQUENCE 601 AA; 64674 MW; 7E7581A11674B4C9 CRC64;

Query Match 3.1%; Score 8; DB 1; Length 601;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 AGDGGGSC 68
|||||
DB 435 AGDGGGSC 442

RESULT 8
GGII STAHA
ID_GGII STAHA STANDARD; PRT; 44 AA.
AC P11697;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Antibacterial protein 1 (Gonococcal growth inhibitor 1).
OS Staphylococcus haemolyticus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1283;
RN [1]
RP SEQUENCE.
RX MEDLINE=88339821; PubMed=3138972;
RA Watson D.C., Yaguchi M., Bisailon J.G., Beaudet R., Morosoli R.;
RT "The amino acid sequence of a gonococcal growth inhibitor from
Staphylococcus haemolyticus."
RL Biochem. J. 252:87-93(1988).
CC -!- DOMAIN: THIS PROTEIN IS POSSIBLY THE SIGNAL SEQUENCE OF A SECRETED
OR MEMBRANE ASSOCIATED PROTEIN.
CC -!- SIMILARITY: HIGH, TO ANTIBACTERIAL PROTEINS 2 AND 3.
DR PIR; S00599; BXSAL.
DR InterPro; IPR008846; Staph_haemo.
DR Pfam; PF05480; Staph_haemo; 1.
KW Antibiotic; Formylation.
FT MOD RES 1 1 FORMYLATION (POTENTIAL).
SQ SEQUENCE 44 AA; 4523 MW; FAB1702DBA3A3238 CRC64;

Query Match 2.7%; Score 7; DB 1; Length 44;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 AAAVSAG 62

Db 9 AAAVSAG 15
|||||
RESULT 9
NULM ALBCO
ID_NULM ALBCO STANDARD; PRT; 99 AA.
AC P48928; 1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 4L (EC 1.6.5.3).
GN ND4L.
OS Albinaria coerulea (Land snail).
OG Mitochondrion.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC Sigmoethra; Clausilioidea; Clausiliidae; Alopiinae; Albinaria.
OX NCBI_TaxID=42349;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96120351; PubMed=7498775;
RA Hatzoglou E., Rodakis G.C., Lecanidou R.;
RT "Complete sequence and gene organization of the mitochondrial genome
of the land snail Albinaria coerulea."
RL Genetics 140:1353-1366(1995).
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
DR EMBL; X83390; CAA58298.1; -.
DR PIR; S59145; S59145.
DR InterPro; IPR001133; Oxidored_4L.
DR Pfam; PF00420; oxidored_g2; 1.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 99 AA; 10909 MW; CE65A360E646AAE5 CRC64;

Query Match 2.7%; Score 7; DB 1; Length 99;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 FAAAAEA 223
|||||
DB 63 FAAAAEA 69

RESULT 10
H33_STYLE
ID_H33_STYLE STANDARD; PRT; 114 AA.
AC P81197;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Histone H3-3 (Fragment).
GN H3-3.
OS Stylonychia lemnae.
OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
OC Stichotrichida; Oxytrichidae; Stylonychia.
OX NCBI_TaxID=5949;
RN [1]
RP SEQUENCE FROM N.A.
RA Bernhard D.;
RT "Several highly divergent histone H3 genes in the hypotrich ciliate
Stylonychia lemnae."
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Histone H3, along with histone H4, plays a central role
in nucleosome formation.
CC -!- SUBUNIT: The nucleosome is an octamer containing two molecules

```

DR Pfam; PF00125; histone; 1.
DR PRINTS; PR00622; HISTONEH3.
DR SMART; SM00428; H3; 1.
DR PROSITE; PS00322; HISTONE_H3_1; FALSE_NEG.
DR PROSITE; PS00959; HISTONE_H3_2; 1.
KW Nuclear protein; Chromosomal protein; DNA-binding; Nucleosome core.
FT NON_TER 1
FT NON_TER 114
SQ SEQUENCE 114 AA; 12794 MW; 0E411EBD51DCEB06 CRC64;

Query Match          2.7%; Score 7; DB 1; Length 114;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AAATGGV 40
   |||||
DB 22 AAATGGV 28

RESULT 12
H37 STYLE STANDARD; PRT; 114 AA.
AC P81201;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Histone H3-7 (Fragment).
GN H3-7.
OS Stylonychia lemnae.
OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
OC Stichotrichida; Oxytrichidae; Stylonychia.
OC NCBI_TaxID=5949;
RN [1]
RP SEQUENCE FROM N.A.
RA Bernhard D.;
RT "Several highly divergent histone H3 genes in the hypotrich ciliate
   stylonychia lemnae.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Histone H3, along with histone H4, plays a central role
   in nucleosome formation.
CC -!- SUBUNIT: The nucleosome is an octamer containing two molecules
   each of H2A, H2B, H3 and H4. The octamer wraps approximately 146
   bp of DNA.
CC -!- SIMILARITY: Belongs to the histone H3 family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; Y16633; CAA76336.1; -.
DR InterPro; IPR007124; Hist_TAF.
DR InterPro; IPR007125; Histone_core_D.
DR InterPro; IPR00164; Histone_H3.
DR Pfam; PF00125; histone; 1.
DR PRINTS; PR00622; HISTONEH3.
DR SMART; SM00428; H3; 1.
DR PROSITE; PS00322; HISTONE_H3_1; FALSE_NEG.
DR PROSITE; PS00959; HISTONE_H3_2; 1.
KW Nuclear protein; Chromosomal protein; DNA-binding; Nucleosome core.
FT NON_TER 1
FT NON_TER 114
SQ SEQUENCE 114 AA; 12689 MW; 36A9B6EA7F1D01FC CRC64;

Query Match          2.7%; Score 7; DB 1; Length 114;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AAATGGV 40
   |||||

```

```
Db          22 AAATGGV 28

RESULT 13
H39_STYLE
ID_H39_STYLE STANDARD; PRT; 114 AA.
AC P81196;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Histone H3-2 (Fragment).
GN H3-2.
OS Stylonychia lemnae.
OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
OC Stichotrichida; Oxytrichidae; Stylonychia.
OX NCBI_TaxID=5949;
RN [1]
RP SEQUENCE FROM N.A.
RA Bernhard D.;
RT "Several highly divergent histone H3 genes in the hypotrich ciliate Stylonychia lemnae.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Histone H3, along with histone H4, plays a central role in nucleosome formation.
CC -!- SUBUNIT: The nucleosome is an octamer containing two molecules each of H2A, H2B, H3 and H4. The octamer wraps approximately 146 bp of DNA.
CC -!- SIMILARITY: Belongs to the histone H3 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y16635; CAA76338.1; -.
DR InterPro; IPR007124; Hist_TAF.
DR InterPro; IPR007125; Histone_core_D.
DR InterPro; IPR000164; Histone_H3.
DR Pfam; PF00125; histone; 1.
DR PRINTS; PR00622; HISTONEH3.
DR PROSITE; PS00322; HISTONE_H3_1; FALSE_NEG.
DR PROSITE; PS00959; HISTONE_H3_2; 1.
KW Nuclear protein; Chromosomal protein; DNA-binding; Nucleosome core.
FT NON_TER 1
FT NON_TER 114
FT NON_TER 114
SQ SEQUENCE 114 AA; 12689 MW; 36A9B6EA7F1D01FC CRC64;

Query Match 2.7%; Score 7; DB 1; Length 114;
Best Local Similarity 100.0%; Pred.No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AAATGGV 40
Db 22 AAATGGV 28

RESULT 14
R18E_SULAC
ID_R18E_SULAC STANDARD; PRT; 118 AA.
AC P39474;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L18e (HL29e).
GN RPL18E.
OS Sulfolobus acidocaldarius.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OX NCBI_TaxID=2285;
RN [1]

Query Match 2.7%; Score 7; DB 1; Length 114;
Best Local Similarity 100.0%; Pred.No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AAATGGV 40
Db 22 AAATGGV 28
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33909 / NCIB 11770 / DSM 639;
RX MEDLINE=95320156; PubMed=7597027;
RA Langer D., Hain J., Thuriaux P., Zillig W.;
RT "Transcription in archaea: similarity to that in eucarya.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:5768-5772(1995).
CC -!- SIMILARITY: Belongs to the L18E family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X80194; CAA56481.1; -.
DR PIR; S47024; S47024.
DR HAMAP; MF_00329; -.
DR InterPro; IPR001196; Ribosomal_L15.
DR InterPro; IPR000039; Ribosomal_L18e.
DR Pfam; PF00256; L15; 1.
DR PROSITE; PS01106; RIBOSOMAL_L18E; 1.
KW Ribosomal protein.
SQ SEQUENCE 118 AA; 13369 MW; 506414A632CB3CF4 CRC64;

Query Match 2.7%; Score 7; DB 1; Length 118;
Best Local Similarity 100.0%; Pred.No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 APRRKRA 50
Db 34 APRRKRA 40

RESULT 15
CRCB_PSEPK
ID_CRCB_PSEPK STANDARD; PRT; 124 AA.
AC Q88FT1;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Protein crCB homolog.
GN CRCB OR PP4001.
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinl C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzes A.,
RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
RA Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the crCB family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
```

```
-----
CC  EMBL; AE016788; AAN69595.1; -.
DR  TIGR; PP4001; -.
DR  HAMAP; MF_00454; -; 1.
DR  InterPro; IPR003691; Camphor_CrcB.
DR  Pfam; PF02537; CRCB; 1.
KW  Transmembrane; Complete proteome.
FT  TRANSMEM 4 21 Potential.
FT  TRANSMEM 34 56 Potential.
FT  TRANSMEM 66 85 Potential.
FT  TRANSMEM 97 119 Potential.
SQ  SEQUENCE 124 AA; 12973 MW; E1A38ED15D0D55A6 CRC64;

Query Match 2.7%; Score 7; DB 1; Length 124;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 AAVSAGG 63
Db 6 AAVSAGG 12
```

Search completed: October 1, 2004, 17:08:32
Job time : 27 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OMprotein - protein search, using sw model

Run on: October 1, 2004, 17:02:22 ; Search time 116 Seconds
(without alignments)
696.316 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 256
Sequence: 1 MGKYMRCRGAAGAEVAAVE.....VRGVPLDAGGRFEWAPVVS I 256

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL 25:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phage:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacteriap:
17: sp_archheap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 10 | 3.9 | 277 | 17 Q9HPI3 | Q9hpi3 halobacteri |
| 2 | 9 | 3.5 | 591 | 16 Q8GP35 | Q8gp35 bordetella |
| 3 | 9 | 3.5 | 664 | 16 Q9RKL3 | Q9rkl3 streptomyce |
| 4 | 8 | 3.1 | 55 | 9 Q853R3 | Q853r3 mycobacteri |
| 5 | 8 | 3.1 | 173 | 10 Q9M2I3 | Q9m2i3 arabidopsis |
| 6 | 8 | 3.1 | 180 | 10 Q8H4B9 | Q8h4b9 oryza sativ |
| 7 | 8 | 3.1 | 181 | 10 Q9M2I6 | Q9m2i6 arabidopsis |
| 8 | 8 | 3.1 | 191 | 16 Q89S92 | Q89s92 bradyrhizob |
| 9 | 8 | 3.1 | 201 | 2 Q8GDM1 | Q8gdm1 heliobacill |
| 10 | 8 | 3.1 | 206 | 16 Q8XWK7 | Q8xwk7 ralstonia s |
| 11 | 8 | 3.1 | 215 | 10 Q9M3U5 | Q9m3u5 nicotiana t |
| 12 | 8 | 3.1 | 215 | 10 Q49911 | Q499i1 nicotiana t |
| 13 | 8 | 3.1 | 221 | 16 Q98ET3 | Q98et3 rhizobium l |
| 14 | 8 | 3.1 | 238 | 16 Q8FYA5 | Q8fya5 brucella su |
| 15 | 8 | 3.1 | 248 | 16 Q89P10 | Q89p10 bradyrhizob |
| 16 | 8 | 3.1 | 250 | 16 Q7U6X9 | Q7u6x9 synechococc |

| | | | | | |
|----|---|-----|-----|-----------|--------------------|
| 17 | 8 | 3.1 | 273 | 11 Q9CVA0 | Q9cva0 mus musculu |
| 18 | 8 | 3.1 | 278 | 2 Q937M8 | Q937m8 photorhabdu |
| 19 | 8 | 3.1 | 279 | 11 Q8K2X7 | Q8k2x7 mus musculu |
| 20 | 8 | 3.1 | 281 | 2 Q8KU07 | Q8ku07 xenophilus |
| 21 | 8 | 3.1 | 295 | 10 Q8GY40 | Q8gy40 arabidopsis |
| 22 | 8 | 3.1 | 320 | 16 Q9EWJ3 | Q9ewj3 streptomyce |
| 23 | 8 | 3.1 | 331 | 16 Q7WDV3 | Q7wdv3 bordetella |
| 24 | 8 | 3.1 | 331 | 16 Q7W2V3 | Q7w2v3 bordetella |
| 25 | 8 | 3.1 | 334 | 16 Q865I3 | Q865i3 streptomyce |
| 26 | 8 | 3.1 | 336 | 16 Q9KZF0 | Q9kzf0 streptomyce |
| 27 | 8 | 3.1 | 353 | 2 Q7WTJ2 | Q7wtj2 acinetobact |
| 28 | 8 | 3.1 | 396 | 12 Q8VA04 | Q8va04 apple stem |
| 29 | 8 | 3.1 | 401 | 8 Q9GCV9 | Q9gcv9 scherffelia |
| 30 | 8 | 3.1 | 404 | 16 Q06263 | Q06263 mycobacteri |
| 31 | 8 | 3.1 | 404 | 16 Q7TWH6 | Q7twh6 mycobacteri |
| 32 | 8 | 3.1 | 410 | 10 Q7XX17 | Q7xxl7 oryza sativ |
| 33 | 8 | 3.1 | 418 | 16 Q8XZX5 | Q8xzx5 ralstonia s |
| 34 | 8 | 3.1 | 427 | 16 Q92BJ2 | Q92bj2 listeria in |
| 35 | 8 | 3.1 | 427 | 16 Q8Y707 | Q8y707 listeria mo |
| 36 | 8 | 3.1 | 469 | 16 Q88AX9 | Q88ax9 pseudomonas |
| 37 | 8 | 3.1 | 491 | 16 Q82B50 | Q82b50 streptomyce |
| 38 | 8 | 3.1 | 492 | 4 Q96T73 | Q96t73 homo sapien |
| 39 | 8 | 3.1 | 493 | 16 Q7UI25 | Q7ui25 rhodopirell |
| 40 | 8 | 3.1 | 504 | 16 Q82T80 | Q82t80 nitrosomona |
| 41 | 8 | 3.1 | 548 | 10 Q9LJF7 | Q9luf7 arabidopsis |
| 42 | 8 | 3.1 | 601 | 5 Q95YF4 | Q95yf4 drosophila |
| 43 | 8 | 3.1 | 618 | 3 Q86ZN5 | Q86zn5 podospora a |
| 44 | 8 | 3.1 | 653 | 10 Q8H8N3 | Q8h8n3 oryza sativ |
| 45 | 8 | 3.1 | 667 | 5 Q9VVFJ6 | Q9vfvj6 drosophila |

ALIGNMENTS

RESULT 1
Q9HPI3
ID Q9HPI3 PRELIMINARY; PRT; 277 AA.
AC Q9HPI3;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Cytochrome aa3 controlling protein.
GN CCP OR VNG1623G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181 (2000).
DR EMBL; AE005072; AAG19884.1; -.
DR PIR; H84314; H84314.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0006461; P:protein complex assembly; IEA.
DR InterPro; IPR003780; COX15 Ctaa.
DR Pfam; PF02628; COX15-Ctaa; 1.
KW Complete proteome.
SQ SEQUENCE 277 AA; 29053 MW; B053123766E274B1 CRC64;

Query Match 3.9%; Score 10; DB 17; Length 277;
Best Local Similarity 100.0%; Pred. No. 0.69;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 32 R5AAATGGVA 41
Db 216 R5AAATGGVA 225

RESULT 2
Q8GP35
ID Q8GP35 PRELIMINARY; PRT; 591 AA.
AC Q8GP35;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Multicopper oxidase (Copper resistance protein).
GN COPA OR BP3315.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22242525; PubMed=12354238;
RA Huston W.M., Jennings M.P., McEwan A.G.;
RT "The multicopper oxidase of Pseudomonas aeruginosa is a ferroxidase
with a central role in iron acquisition.";
RL Mol. Microbiol. 45:1741-1750(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; AF455754; AAN52533.1; -.
DR EMBL; BX640421; CAE43580.1; -.
DR GO; GO:0005743; C:mitochondrial inner membrane; IEA.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR006376; CopA.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR002355; Cu_ox_copper_BS.
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR006311; Tat.
DR Pfam; PF00394; Cu-oxidase; 3.
DR TIGRFAMs; TIGR01480; copper_res_A; 1.
DR TIGRFAMs; TIGR01409; TAT_signal_seq; 1.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 2.
DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
KW Complete proteome.
SQ SEQUENCE 591 AA; 65643 MW; BFA5587C557DE393 CRC64;

Query Match 3.5%; Score 9; DB 16; Length 591;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 GAAGQGGAA 108
Db 384 GAAGQGGAA 392

RESULT 3
```

```
Q9RKL3
ID Q9RKL3 PRELIMINARY; PRT; 664 AA.
AC Q9RKL3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative integral membrane protein.
GN SCO4069 OR SCD25.05.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939118; CAB56349.1; -.
KW Complete proteome.
SQ SEQUENCE 664 AA; 69158 MW; C3F83EC120526F2A CRC64;

Query Match 3.5%; Score 9; DB 16; Length 664;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 GAALAAAGLS 114
Db 527 GAALAAAGLS 535

RESULT 4
Q853R3
ID Q853R3 PRELIMINARY; PRT; 55 AA.
AC Q853R3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp14.
OS Mycobacteriophage Bx1.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=205877;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22592660; PubMed=12705866;
RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,
RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,
RA Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,
RA Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,
RA Hatfull G.F.;
RT "Origins of highly mosaic mycobacteriophage genomes.";
RL Cell 113:171-182(2003).
DR EMBL; AY129337; AAN16674.1; -.
SQ SEQUENCE 55 AA; 5697 MW; E4EBAC5A9B3AC3EB CRC64;

Query Match 3.1%; Score 8; DB 9; Length 55;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 AALAAAGLS 114
Db 48 AALAAAGLS 55
```

RESULT 5

Q9M2I3 Q9M2I3 PRELIMINARY; PRT; 173 AA.
AC Q9M2I3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN F9D24.240.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA D'Angelo M., Vezzi A., Modesto D., Pigazzi M., Valle G., Mewes H.W.,
RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL137081; CAB68171.1; -.
DR PIR; T45993; T45993.
KW Hypothetical protein.
SQ SEQUENCE 173 AA; 19724 MW; ACD2BF15B48B0C2D CRC64;

Query Match 3.1%; Score 8; DB 10; Length 173;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 LSDLES DL 187
|||||
Db 148 LSDLES DL 155

RESULT 6

Q8H4B9 Q8H4B9 PRELIMINARY; PRT; 180 AA.
AC Q8H4B9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative translational inhibitor protein.
GN P0048D08.27.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
RT clone: P0048D08.27";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004269; BAC20708.1; -.
DR InterPro; IPR006175; Endoribon LPSP.
DR InterPro; IPR006056; Yj9F-like.
DR Pfam; PF01042; ribonuc L-PSI; 1.
DR TIGRFAMs; TIGR00004; TIGR00004; 1.
DR PROSITE; PS01094; UPF0076; 1.
SQ SEQUENCE 180 AA; 18818 MW; BCE6915086D91B44 CRC64;

Query Match 3.1%; Score 8; DB 10; Length 180;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 198 AATPAE L 205
|||||

Db 14 AATPAE L 21

RESULT 7

Q9M2I6 Q9M2I6 PRELIMINARY; PRT; 181 AA.
AC Q9M2I6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN F9D24.210.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA D'Angelo M., Vezzi A., Modesto D., Pigazzi M., Valle G., Mewes H.W.,
RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL137081; CAB68168.1; -.
DR PIR; T45990; T45990.
KW Hypothetical protein.
SQ SEQUENCE 181 AA; 20661 MW; F8EA87E08FBA48B0 CRC64;

Query Match 3.1%; Score 8; DB 10; Length 181;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 LSDLES DL 187
|||||
Db 154 LSDLES DL 161

RESULT 8

Q89S92 Q89S92 PRELIMINARY; PRT; 191 AA.
AC Q89S92;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE BL12513 protein.
GN BL12513.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005944; BAC47778.1; -.
DR InterPro; IPR001173; Glyco_trans_2.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00535; Glycos_transf_2; 1.
DR PROSITE; PS00038; HLH_1; 1.
KW Complete proteome.
SQ SEQUENCE 191 AA; 20410 MW; 3E1A04A0E09A0413 CRC64;

Query Match 3.1%; Score 8; DB 16; Length 191;

Best Local Similarity 100.0%; Pred. No. 39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 QGAALAAAG 112
Db 64 QGAALAAAG 71

RESULT 9

Q8GDV1 PRELIMINARY; PRT; 201 AA.
AC Q8GDV1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cobinamide kinase (EC 2.7.1.-) (Fragment).
OS Helicobacterium mobilis.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Helicobacteriaceae;
OC Helicobacterillus.
OX NCBI_TaxID=28064;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22337798; PubMed=12446909;
RA Raymond J., Zhaxybayeva O., Gogarten J.P., Gerdes S.Y.,
RA Blankenship R.E.;
RT "Whole-genome analysis of photosynthetic prokaryotes.";
RL Science 298:1616-1620(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Liolios K.G., Chu L., Ostrovskaya O., Mendybaeva N., Koukharenko V.,
RA Gerdes S., Kyrpides N., Overbeek R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY142871; AAN87475.1; -.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR003203; CobU.
DR Pfam; PF02283; CobU; 1.
KW Kinase; Transferase.
FT NON TER 201 201
SQ SEQUENCE 201 AA; 22473 MW; 48D6DC533C37A759 CRC64;

Query Match 3.1%; Score 8; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 AAEAGGDH 148
Db 27 AAEAGGDH 34

RESULT 10

Q8XWK7 PRELIMINARY; PRT; 206 AA.
AC Q8XWK7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Probable transmembrane protein.
GN RSC2467 OR RS01138.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cumnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;

RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL; AL646070; CAD16174.1; -.
KW Complete proteome.
SQ SEQUENCE 206 AA; 21783 MW; 8D49E45A150B8440 CRC64;

Query Match 3.1%; Score 8; DB 16; Length 206;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 AALAAGLS 114
Db 5 AALAAGLS 12

RESULT 11

Q9M3U5 PRELIMINARY; PRT; 215 AA.
AC Q9M3U5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DREPP2 protein.
GN DREPP2.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RA Bassac G., Domergue O., Grignon N., Ferriere N., Escoute J.,
RA Teyssendier de la Serve B.;
RT "Cloning promoters of two developmentally regulated tobacco genes.
RT Study of their spatio-temporal expression pattern.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ277898; CAB91552.1; -.
DR InterPro; IPR008469; DREPP.
DR Pfam; PF05558; DREPP; 1.
SQ SEQUENCE 215 AA; 22915 MW; 2AD0E2F7DEF0346E CRC64;

Query Match 3.1%; Score 8; DB 10; Length 215;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 PAATPAAE 204
Db 176 PAATPAAE 183

RESULT 12

O49911 PRELIMINARY; PRT; 215 AA.
AC O49911;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasma membrane polypeptide.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Logan D.C., Domergue O., Teyssendier de la Serve B., Rossignol M.;
RT "A new family of plasma membrane intrinsic polypeptides differentially
RT regulated during plant development.";
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y08609; CAA69901.1; -.
DR PIR; T03680; T03680.
DR InterPro; IPR008469; DREPP.

```
DR Pfam; PF05558; DREPP; 1.
SQ SEQUENCE 215 AA; 22943 MW; 9DC7E2F7DEB909D8 CRC64;

Query Match      3.1%; Score 8; DB 10; Length 215;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 PAATPAAE 204
Db 176 PAATPAAE 183

RESULT 13
Q98ET3 PRELIMINARY; PRT; 221 AA.
AC Q98ET3;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein mlr4094.
GN MLR4094.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003003; BAB50834.1; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR006402; HAD-SF-IA-v3.
DR Pfam; PF00702; Hydrolase; 1.
DR TIGRFAMs; TIGR01509; HAD-SF-IA-v3; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 221 AA; 23018 MW; 477A3DC6B96F6F29 CRC64;

Query Match      3.1%; Score 8; DB 16; Length 221;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 AAEEAAQA 225
Db 169 AAEEAAQA 176

RESULT 14
Q8FYA5 PRELIMINARY; PRT; 238 AA.
AC Q8FYA5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Thiol:disulfide interchange protein, putative.
GN BR1980.
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1330 / Biovar 1;
RX MEDLINE=22247741; PubMed=12271122;
```

```
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J., Van Aken S.E.,
RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts."
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
DR EMBL; AE014487; AAN30870.1; -.
DR TIGR; BR1980; -.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR006662; ThioRed.
DR InterPro; IPR006663; ThioRedox dom2.
DR PROSITE; PS00194; THIOREDOXIN; 1.
KW Complete proteome.
SQ SEQUENCE 238 AA; 24862 MW; F9BABB6EC07B7C58 CRC64;

Query Match      3.1%; Score 8; DB 16; Length 238;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AAATGGVA 41
Db 79 AAATGGVA 86

RESULT 15
Q89P10 PRELIMINARY; PRT; 248 AA.
AC Q89P10;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Oxidoreductase.
GN BLR3673.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110."
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005948; BAC48938.1; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; ADH_short.
DR Pfam; PF00106; adh short; 1.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; 1.
KW Complete proteome.
SQ SEQUENCE 248 AA; 25207 MW; D137219FCA7FC0C6 CRC64;

Query Match      3.1%; Score 8; DB 16; Length 248;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 GAALAAAGL 113
Db 42 GAALAAAGL 49

Search completed: October 1, 2004, 17:10:35
Job time : 119 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 16:54:37 ; Search time 125 seconds
(without alignments)
578.657 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 256
Sequence: 1 MGKYMRCRGAAGAEVAE.....VRGVPLDAGGRFEWAPVVISI 256

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 256 | 100.0 | 256 | 6 | ABB98757 Maize CKI |
| 2 | 101 | 39.5 | 125 | 3 | AAB26250 Corn cycl |
| 3 | 42 | 16.4 | 42 | 6 | ABB98860 Maize CKI |
| 4 | 21 | 8.2 | 46 | 6 | ABB98862 Maize CKI |
| 5 | 21 | 8.2 | 53 | 3 | AAB26245 Corn cycl |
| 6 | 21 | 8.2 | 53 | 3 | AAP01940 Corn cycl |
| 7 | 21 | 8.2 | 116 | 5 | ABG65674 Corn ICK |
| 8 | 21 | 8.2 | 126 | 3 | AAB26249 Corn cycl |
| 9 | 17 | 6.6 | 227 | 6 | ABB98759 Maize CKI |
| 10 | 17 | 6.6 | 262 | 5 | ABG65670 OSICK 2 p |
| 11 | 10 | 3.9 | 10 | 5 | ABG65741 Plant ICK |
| 12 | 10 | 3.9 | 226 | 5 | ABG65692 Rice OsIC |
| 13 | 9 | 3.5 | 25 | 5 | ABG65793 Plant pot |
| 14 | 9 | 3.5 | 37 | 3 | AAP01950 Rice cycl |
| 15 | 9 | 3.5 | 37 | 3 | AAB27252 Rice cycl |
| 16 | 9 | 3.5 | 324 | 6 | ABP99213 Orthosomy |
| 17 | 9 | 3.5 | 580 | 6 | AAO19831 Bordetell |
| 18 | 8 | 3.1 | 8 | 5 | ABG65707 Plant ICK |
| 19 | 8 | 3.1 | 8 | 5 | ABG65767 Plant ICK |
| 20 | 8 | 3.1 | 45 | 3 | AAB52025 Human sec |
| 21 | 8 | 3.1 | 62 | 3 | AAY88379 Partial t |
| 22 | 8 | 3.1 | 62 | 3 | AAB36150 Mosquito |
| 23 | 8 | 3.1 | 62 | 3 | AAB37950 Trypsin m |
| 24 | 8 | 3.1 | 62 | 3 | AAB28345 Trypsin M |
| 25 | 8 | 3.1 | 62 | 7 | ADC35463 Trypsin m |

| | | | | | | |
|----|---|-----|-----|---|----------|--------------------|
| 26 | 8 | 3.1 | 62 | 7 | ADD10243 | ADD10243 Mosquito |
| 27 | 8 | 3.1 | 209 | 4 | AAM01295 | AAM01295 Human tra |
| 28 | 8 | 3.1 | 209 | 4 | AAU69940 | AAU69940 Human tra |
| 29 | 8 | 3.1 | 209 | 4 | ABU71831 | ABU71831 Prostate |
| 30 | 8 | 3.1 | 209 | 5 | ABB95400 | ABB95400 Human tra |
| 31 | 8 | 3.1 | 209 | 6 | ABR54512 | ABR54512 Partial H |
| 32 | 8 | 3.1 | 209 | 7 | ADB14347 | ADB14347 Human tra |
| 33 | 8 | 3.1 | 237 | 4 | ABB69862 | ABB69862 Drosophil |
| 34 | 8 | 3.1 | 243 | 6 | ABM67642 | ABM67642 Photorhab |
| 35 | 8 | 3.1 | 336 | 6 | ABU27612 | ABU27612 Protein e |
| 36 | 8 | 3.1 | 404 | 7 | ADB80207 | ADB80207 Mycobacte |
| 37 | 8 | 3.1 | 427 | 5 | ABB49079 | ABB49079 Listeria |
| 38 | 8 | 3.1 | 427 | 6 | ABU32600 | ABU32600 Protein e |
| 39 | 8 | 3.1 | 492 | 3 | AAY44406 | AAY44406 Human 20P |
| 40 | 8 | 3.1 | 492 | 3 | AAY77726 | AAY77726 Human tum |
| 41 | 8 | 3.1 | 492 | 3 | AAY57280 | AAY57280 Ovr115 ho |
| 42 | 8 | 3.1 | 492 | 3 | AAY92050 | AAY92050 HrPCa6/7 |
| 43 | 8 | 3.1 | 492 | 3 | AAB36901 | AAB36901 Human TMP |
| 44 | 8 | 3.1 | 492 | 4 | AAM01294 | AAM01294 Human tra |
| 45 | 8 | 3.1 | 492 | 4 | AAM01315 | AAM01315 P1000C am |

ALIGNMENTS

RESULT 1
ABB98757
ID ABB98757 standard; protein; 256 AA.
XX
AC ABB98757;
XX
DT 21-FEB-2003 (first entry)
XX
DE Maize CKI_B.
XX
KW Maize; cyclin-dependent kinase inhibitor; CKI; CKI B; plant; corn;
KW crop yield; root size; plant growth; tassel size; ear size;
KW male sterility; endoreduplication.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT Domain 1..7
FT /note= "Conserved domain"
XX
PN WO200281623-A2.
XX
PD 17-OCT-2002.
XX
PF 06-NOV-2001; 2001WO-US044038.
XX
PR 07-NOV-2000; 2000US-0246349P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
PA (ARIZ-) ARIZONA BOARD OF REGENTS.
XX
PI Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;
XX
DR WPI; 2003-058511/05.
DR N-PSDB; ABV74603.
XX
PT Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
PT involved in cell cycle regulation, and useful for altering cell cycle
PT protein content, cell cycle progression, cell number and composition of
PT plants.
XX
PS Claim 12; Page 65-66; 69pp; English.
XX
CC The present sequence is maize cyclin-dependent kinase inhibitor (CKI),
CC CKI_B. The coding sequence for this protein (I) is useful for modulating
CC the activity of cyclin-dependent kinase (CDK) in a plant such as a corn,
CC soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice,
CC barley, oil-seed Brassica and millet. Modulating the activity of CDK,

CC preferably modulating downward is useful for providing differential
CC growth in a plant, especially a positive growth advantage and modulating
CC CDK activity upward is useful for increasing crop yield, root size, plant
CC growth, tassel size and/or ear size. Modulating CDK activity is also
CC useful for conferring male sterility and for improving transformation
CC frequencies by increasing the number of cells in cell division. CDK
CC activity can also be modulated for modulating endoreduplication in the
CC endosperm of corn, sorghum, wheat, rice, barley, and millet, where the
CC promoter used is an endosperm-preferred promoter. The cell numbers are
CC modulated in one or more tissues of a plant, comprising root, seed,
CC tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves,
CC stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm,
CC pericarp, meristematic or leaf cells. (I) is also useful for identifying
CC maize CKI interacting proteins, by adducting the nucleic acid sequence to
CC a second nucleic acid sequence encoding a DNA-binding domain
XX
SQ Sequence 256 AA;

Query Match 100.0%; Score 256; DB 6; Length 256;
Best Local Similarity 100.0%; Pred. No. 1.7e-234;
Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGKYMRCGGAAGAEVAAVEVTQVGVTRSRSAATGGVAKVAPRRKRAPAGEPAAAVS 60
Db 1 MGKYMRCGGAAGAEVAAVEVTQVGVTRSRSAATGGVAKVAPRRKRAPAGEPAAAVS 60
QY 61 AGDGGGSCYIHLRSRMLFMAPPQPPQPSVDSVPTPVEAADGAAGQQAALAAAGLRCSSSTA 120
Db 61 AGDGGGSCYIHLRSRMLFMAPPQPPQPSVDSVPTPVEAADGAAGQQAALAAAGLRCSSSTA 120
QY 121 SSVNLGLGQSGSHTCRSYDAAEAGGHHVLDVSAASNSGSGPDRRERTTPSSRAHGEL 180
Db 121 SSVNLGLGQSGSHTCRSYDAAEAGGHHVLDVSAASNSGSGPDRRERTTPSSRAHGEL 180
QY 181 SDLESIDLGHKTGSPSLPAATPAAELIVPPAHEIQEFAAAEAAQAKRFASKYNFDFVRGV 240
Db 181 SDLESIDLGHKTGSPSLPAATPAAELIVPPAHEIQEFAAAEAAQAKRFASKYNFDFVRGV 240
QY 241 PLDAGGRFEWAPVVS I 256
Db 241 PLDAGGRFEWAPVVS I 256

RESULT 2
AAB26250
ID AAB26250 standard; protein; 125 AA.
XX
AC AAB26250;
XX
DT 17-JAN-2001 (first entry)
XX
DE Corn cyclin-dependent kinase inhibitor #3.
XX
KW Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
KW cell growth; herbicide.
XX
OS Zea mays.
XX
PN WO200060087-A2.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US009106.
XX
FF
XX
PR 07-APR-1999; 99US-0128192P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Klein TM, Weng Z, Cahoon RE;
XX
DR WPI; 2000-679375/66.
DR N-PSDB; AAA95281.
XX

PT Cyclin dependent kinase inhibitor sequences, useful for identifying
PT herbicides and plant growth inhibitors.
XX
PS Claim 10; Fig 1; 58pp; English.
XX
CC The present sequence is the corn cyclin-dependent kinase inhibitor
CC (CDKI). Its coding sequence was isolated by searching a contig comprising
CC cDNA from corn roots and ear leaf sheath for sequences similar to those
CC encoding the CDKI from *Chenopodium rubrum*, *Caenorhabditis elegans* and
CC *Arabidopsis thaliana*. CDKI is involved in the cell cycle, and may promote
CC or inhibit cell division and growth. The protein and its coding sequence
CC are useful in the production of transgenic plants which produce increased
CC or decreased amounts of the CDKI protein, in the identification of
CC herbicides, in genetic and physical mapping and in the isolation of the
CC CDKI gene in other organisms
XX
SQ Sequence 125 AA;

Query Match 39.5%; Score 101; DB 3; Length 125;
Best Local Similarity 100.0%; Pred. No. 1.4e-87;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 156 ASNSGSGPDRRERTTPSSRAHGELSDLESIDLGHKTGSPSLPAATPAAELIVPPAHEIQE 215
Db 25 ASNSGSGPDRRERTTPSSRAHGELSDLESIDLGHKTGSPSLPAATPAAELIVPPAHEIQE 84
QY 216 FFAAAEAAQAKRFASKYNFDFVRGVPLDAGGRFEWAPVVS I 256
Db 85 FFAAAEAAQAKRFASKYNFDFVRGVPLDAGGRFEWAPVVS I 125

RESULT 3
AAB98860
ID AAB98860 standard; protein; 42 AA.
XX
AC ABB98860;
XX
DT 21-FEB-2003 (first entry)
XX
DE Maize CKI_B C-terminal protein fragment.
XX
KW Maize; cyclin-dependent kinase inhibitor; CKI; plant; corn; crop yield;
KW root size; plant growth; tassel size; ear size; male sterility;
KW endoreduplication.
XX
OS Zea mays.
XX
PN WO200281623-A2.
XX
PD 17-OCT-2002.
XX
PF 06-NOV-2001; 2001WO-US044038.
XX
PR 07-NOV-2000; 2000US-0246349P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
PA (ARIZ-) ARIZONA BOARD OF REGENTS.
XX
PI Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;
XX
DR WPI; 2003-058511/05.
XX
PT Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
PT involved in cell cycle regulation, and useful for altering cell cycle
PT protein content, cell cycle progression, cell number and composition of
PT plants.
XX
PS Example 11; Page 46; 69pp; English.
XX
CC The present invention relates to maize cyclin-dependent kinase inhibitor
CC (CKI), CKI_B, CKI_C and CKI_D. The coding sequences (I) for these
CC proteins are useful for modulating the activity of CDK in a plant such as
CC a corn, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton,

CC rice, barley, oil-seed Brassica and millet. Modulating the activity of
 CC CDK, preferably modulating downward is useful for providing differential
 CC growth in a plant, especially a positive growth advantage and modulating
 CC CDK activity upward is useful for increasing crop yield, root size, plant
 CC growth, tassel size and/or ear size. Modulating CDK activity is also
 CC useful for conferring male sterility and for improving transformation
 CC frequencies by increasing the number of cells in cell division. CDK
 CC activity can also be modulated for modulating endoreduplication in the
 CC endosperm of corn, sorghum, wheat, rice, barley, and millet, where the
 CC promoter used is an endosperm-preferred promoter. The cell numbers are
 CC modulated in one or more tissues of a plant, comprising root, seed,
 CC tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves,
 CC stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm,
 CC pericarp, meristematic or leaf cells. (I) is also useful for identifying
 CC maize CKI interacting proteins, by adducting the nucleic acid sequence to
 CC a second nucleic acid sequence encoding a DNA-binding domain. The present
 CC sequence is a C-terminal protein fragment of CKI_B. This sequence
 CC contains a CDK binding region and/or cyclin binding domains
 XX

SQ Sequence 42 AA;

Query Match 16.4%; Score 42; DB 6; Length 42;
 Best Local Similarity 100.0%; Pred. No. 5.3e-32;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 EFAAAEAQAQKRFASKYNFDFVRGVPLDAGGRFEWAPVVS I 256
 Db 1 EFAAAEAQAQKRFASKYNFDFVRGVPLDAGGRFEWAPVVS I 42

RESULT 4

ABB98862
 ID ABB98862 standard; protein; 46 AA.

AC ABB98862;

DT 21-FEB-2003 (first entry)

DE Maize CKI_D C-terminal protein fragment.

XX Maize; cyclin-dependent kinase inhibitor; CKI; plant; corn; crop yield;
 KW root size; plant growth; tassel size; ear size; male sterility;
 KW endoreduplication.

OS Zea mays.

PN WO200281623-A2.

PD 17-OCT-2002.

PF 06-NOV-2001; 2001WO-US044038.

PR 07-NOV-2000; 2000US-0246349P.

XX (PION-) PIONEER HI-BRED INT INC.
 PA (ARIZ-) ARIZONA BOARD OF REGENTS.

PI Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;

XX WPI; 2003-058511/05.

PT Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
 PT involved in cell cycle regulation, and useful for altering cell cycle
 PT protein content, cell cycle progression, cell number and composition of
 PT plants.

PS Example 11; Page 46; 69pp; English.

XX The present invention relates to maize cyclin-dependent kinase inhibitor
 CC (CKI), CKI_B, CKI_C and CKI_D. The coding sequences (I) for these
 CC proteins are useful for modulating the activity of CDK in a plant such as
 CC a corn, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton,
 CC rice, barley, oil-seed Brassica and millet. Modulating the activity of

CC CDK, preferably modulating downward is useful for providing differential
 CC growth in a plant, especially a positive growth advantage and modulating
 CC CDK activity upward is useful for increasing crop yield, root size, plant
 CC growth, tassel size and/or ear size. Modulating CDK activity is also
 CC useful for conferring male sterility and for improving transformation
 CC frequencies by increasing the number of cells in cell division. CDK
 CC activity can also be modulated for modulating endoreduplication in the
 CC endosperm of corn, sorghum, wheat, rice, barley, and millet, where the
 CC promoter used is an endosperm-preferred promoter. The cell numbers are
 CC modulated in one or more tissues of a plant, comprising root, seed,
 CC tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves,
 CC stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm,
 CC pericarp, meristematic or leaf cells. (I) is also useful for identifying
 CC maize CKI interacting proteins, by adducting the nucleic acid sequence to
 CC a second nucleic acid sequence encoding a DNA-binding domain. The present
 CC sequence is a C-terminal protein fragment of CKI_D. This sequence
 CC contains a CDK binding region and/or cyclin binding domains
 XX

SQ Sequence 46 AA;

Query Match 8.2%; Score 21; DB 6; Length 46;
 Best Local Similarity 100.0%; Pred. No. 5.1e-12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 AKRFASKYNFDFVRGVPLDAG 245
 Db 16 AKRFASKYNFDFVRGVPLDAG 36

RESULT 5

AAB26245

ID AAB26245 standard; protein; 53 AA.

AC AAB26245;

DT 17-JAN-2001 (first entry)

DE Corn cyclin-dependent kinase inhibitor #1.

XX Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
 KW cell growth; herbicide.

OS Zea mays.

PN WO200060087-A2.

PD 12-OCT-2000.

PF 06-APR-2000; 2000WO-US009106.

PR 07-APR-1999; 99US-0128192P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

PA Klein TM, Weng Z, Cahoon RE;

PI WPI; 2000-679375/66.

XX N-PSDB; AAA95276.

PT Cyclin dependent kinase inhibitor sequences, useful for identifying
 PT herbicides and plant growth inhibitors.

PS Claim 10; Page 40; 58pp; English.

XX The present sequence is the corn cyclin-dependent kinase inhibitor
 CC (CDKI). Its coding sequence was isolated by searching a corn silk cDNA
 CC library for sequences similar to those encoding the CDKI from Chenopodium
 CC rubrum, Caenorhabditis elegans and Arabidopsis thaliana. CDKI is involved
 CC in the cell cycle, and may promote or inhibit cell division and growth.
 CC The protein and its coding sequence are useful in the production of
 CC transgenic plants which produce increased or decreased amounts of the
 CC CDKI protein, in the identification of herbicides, in genetic and
 CC physical mapping and in the isolation of the CDKI gene in other organisms

```
XX SQ Sequence 53 AA;
      Query Match      8.2%; Score 21; DB 3; Length 53;
      Best Local Similarity 100.0%; Pred. No. 5.8e-12;
      Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 AKRFASKYNFDFVRGVPLDAG 245
Db 27 AKRFASKYNFDFVRGVPLDAG 47

RESULT 6
AAP01940
ID AAP01940 standard; protein; 53 AA.
XX
AC AAP01940;
XX
DT 01-NOV-2001 (first entry)
XX
DE Corn Cyclin dependent kinase inhibitor (CDKI) clone csiln.pk0050.e6.
XX
KW Cyclin dependent kinase inhibitor; CDKI; herbicide; cell cycle; corn;
  plant growth inhibitor.
XX
OS Zea mays.
XX
PN WO200060087-A2.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US009106.
XX
PR 07-APR-1999; 99US-0128192P.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI Klein TM, Weng Z, Cahoon RE;
XX
DR WPI; 2000-679375/66.
XX
DR N-PSDB; AAN02390.
XX
PT Cyclin dependent kinase inhibitor sequences, useful for identifying
  herbicides and plant growth inhibitors.
XX
PS Claim 10; Page 40; 58pp; English.
XX
CC The invention describes a novel isolated polynucleotide comprising a
  nucleotide sequence encoding one of 17 specific cyclin dependent kinase
  inhibitor (CDKI) polypeptides, cell cycle regulators involved in control
  of cell division, growth and death. The nucleotide sequences can be used
  in a vector to transform a host cell to produce the CDKI polypeptide.
  They can also be used in methods for selecting and obtaining a nucleic
  acid sequence that encodes CDKI or affects the level of CDKI expression.
  The encoded protein can be used in a method for evaluating a compound for
  its ability to inhibit the activity of a CDKI. The inhibitors can be used
  as herbicides. They can also be used to inhibit plant growth. The
  polynucleotide sequences can be used in gene mapping and as genetic
  markers. The sequence is the corn CDKI clone csiln.pk0050.e6 as described
  in the method of the invention
XX
SQ Sequence 53 AA;
      Query Match      8.2%; Score 21; DB 3; Length 53;
      Best Local Similarity 100.0%; Pred. No. 5.8e-12;
      Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 AKRFASKYNFDFVRGVPLDAG 245
Db 27 AKRFASKYNFDFVRGVPLDAG 47

RESULT 7
```

```
ABG65674
ID ABG65674 standard; protein; 116 AA.
XX
AC ABG65674;
XX
DT 07-AUG-2003 (revised)
DT 27-AUG-2002 (first entry)
XX
DE Corn ICK 1 protein.
XX
KW Plant; inhibitor of cyclin dependent kinase; ICK.
XX
OS Zea mays.
XX
PN WO200228893-A2.
XX
PD 11-APR-2002.
XX
PF 29-JUN-2001; 2001WO-IB001492.
XX
PR 14-JUL-2000; 2000US-0218471P.
PR 13-OCT-2000; 2000US-0241219P.
XX
PA (CROP-) CROPDESIGN NV.
XX
PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
  Hatzfeld Y;
XX
DR WPI; 2002-471311/50.
DR N-PSDB; ABK93954.
XX
PT Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
  to screen substrates, drugs or compounds which modulate ICK activity and
  treat disorders characterized by an insufficient or excessive production
  of ICK inhibitors.
XX
PS Claim 48; Fig 4; 141pp; English.
XX
CC This invention relates to the DNA and protein sequences of novel isolated
  ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
  the invention may be used for treating disorders characterised by
  insufficient or excessive production of an ICK inhibitor. The protein of
  the invention may also be used to screen for naturally-occurring ICK
  substrates, drugs or compounds which modulate ICK activity, as well as to
  treat disorders characterised by insufficient or excessive production of
  ICK protein, forms which have decreased or aberrant activity compared to
  ICK wild type protein. The present sequence represents an inhibitor of
  cyclin dependent kinase (ICK) protein of the invention. (Updated on 07-
  AUG-2003 to correct OS field.)
XX
SQ Sequence 116 AA;
      Query Match      8.2%; Score 21; DB 5; Length 116;
      Best Local Similarity 100.0%; Pred. No. 1.2e-11;
      Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 AKRFASKYNFDFVRGVPLDAG 245
Db 86 AKRFASKYNFDFVRGVPLDAG 106

RESULT 8
AAB26249
ID AAB26249 standard; protein; 126 AA.
XX
AC AAB26249;
XX
DT 17-JAN-2001 (first entry)
XX
DE Corn cyclin-dependent kinase inhibitor #2.
XX
KW Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
  cell growth; herbicide.
```


XX OS Zea mays.
XX PN WO2000060087-A2.
XX PD 12-OCT-2000.
XX PF 06-APR-2000; 2000WO-US0009106.
XX PR 07-APR-1999; 99US-0128192P.
XX PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX PI Klein TM, Weng Z, Cahoon RE;
XX WPI; 2000-679375/66.
XX N-PSDB; AAA95280.
XX Cyclin dependent kinase inhibitor sequences, useful for identifying herbicides and plant growth inhibitors.
XX Claim 10; Fig 1; 58pp; English.
XX The present sequence is the corn cyclin-dependent kinase inhibitor (CDKI). Its coding sequence was isolated by searching a corn silk cDNA library for sequences similar to those encoding the CDKI from Chenopodium rubrum, Caenorhabditis elegans and Arabidopsis thaliana. CDKI is involved in the cell cycle, and may promote or inhibit cell division and growth. The protein and its coding sequence are useful in the production of transgenic plants which produce increased or decreased amounts of the CDKI protein, in the identification of herbicides, in genetic and physical mapping and in the isolation of the CDKI gene in other organisms
XX Sequence 126 AA;
SQ
Query Match 8.2%; Score 21; DB 3; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.3e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 225 AKRFASKYNFDFVRGVPLDAG 245
DB 96 AKRFASKYNFDFVRGVPLDAG 116
|||||
RESULT 9
ABB98759
ID ABB98759 standard; protein; 227 AA.
AC ABB98759;
XX 21-FEB-2003 (first entry)
DT Maize CKI_D.
DE
XX Maize; cyclin-dependent kinase inhibitor; CKI; CKI_D; plant; corn; crop yield; root size; plant growth; tassel size; ear size; male sterility; endoreduplication.
KW
KW Zea mays.
XX
XX Key Location/Qualifiers
FH Domain 1..7
FT /note= "Conserved domain"
FT Misc-difference 111
FT /note= "Encoded by CAR"
XX
XX WO200281623-A2.
XX
XX 17-OCT-2002.
PD
XX
XX 06-NOV-2001; 2001WO-US044038.
PF
XX
XX 07-NOV-2000; 2000US-0246349P.
PR

XX (PION-) PIONEER HI-BRED INT INC.
PA (ARIZ-) ARIZONA BOARD OF REGENTS.
XX
XX Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkens BR, Sun Y;
XX WPI; 2003-058511/05.
DR N-PSDB; ABV74605.
XX
XX Novel cyclin-dependent kinase polynucleotides and their encoded proteins, involved in cell cycle regulation, and useful for altering cell cycle protein content, cell cycle progression, cell number and composition of plants.
XX Claim 12; Page 69; 69pp; English.
XX The present sequence is maize cyclin-dependent kinase inhibitor (CKI), CKI D. The coding sequence for this protein (I) is useful for modulating the activity of cyclin-dependent kinase (CDK) in a plant such as a corn, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley, oil-seed Brassica and millet. Modulating the activity of CDK, preferably modulating downward is useful for providing differential growth in a plant, especially a positive growth advantage and modulating CDK activity upward is useful for increasing crop yield, root size, plant growth, tassel size and/or ear size. Modulating CDK activity is also useful for conferring male sterility and for improving transformation frequencies by increasing the number of cells in cell division. CDK activity can also be modulated for modulating endoreduplication in the endosperm of corn, sorghum, wheat, rice, barley, and millet, where the promoter used is an endosperm-preferred promoter. The cell numbers are modulated in one or more tissues of a plant, comprising root, seed, tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves, stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm, pericarp, meristematic or leaf cells. (I) is also useful for identifying maize CKI interacting proteins, by adducting the nucleic acid sequence to a second nucleic acid sequence encoding a DNA-binding domain
XX Sequence 227 AA;
SQ
Query Match 6.6%; Score 17; DB 6; Length 227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 23 QVGVGVRTRSRSAATGG 39
DB 23 QVGVGVRTRSRSAATGG 39
|||||
RESULT 10
ABG5670
ID ABG5670 standard; protein; 262 AA.
XX
XX AC ABG5670;
XX
XX 27-AUG-2002 (first entry)
DT
XX
XX OsICK 2 protein.
DE
XX
XX Plant; inhibitor of cyclin dependent kinase; ICK.
XX
XX Oryza sativa.
OS
XX
XX WO200228893-A2.
XX
XX 11-APR-2002.
PD
XX
XX 29-JUN-2001; 2001WO-IB001492.
XX
XX 14-JUL-2000; 2000US-0218471P.
XX
XX 13-OCT-2000; 2000US-0241219P.
XX
XX (CROP-) CROPDESIGN NV.
PA
XX

PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
PI Hatzfeld Y;
XX
DR WPI; 2002-471311/50.
DR N-PSDB; ABK93958.
XX
PT Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
PT to screen substrates, drugs or compounds which modulate ICK activity and
PT treat disorders characterized by an insufficient or excessive production
PT of ICK inhibitors.
XX
PS Claim 48; Fig 3; 141pp; English.
XX
CC This invention relates to the DNA and protein sequences of novel isolated
CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
CC the invention may be used for treating disorders characterised by
CC insufficient or excessive production of an ICK inhibitor. The protein of
CC the invention may also be used to screen for naturally-occurring ICK
CC substrates, drugs or compounds which modulate ICK activity, as well as to
CC treat disorders characterised by insufficient or excessive production of
CC ICK protein, forms which have decreased or aberrant activity compared to
CC ICK wild type protein. The present sequence represents an inhibitor of
CC cyclin dependent kinase (ICK) protein of the invention
XX
SQ Sequence 262 AA;

Query Match 6.6%; Score 17; DB 5; Length 262;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 EVTQVVGVTRSRSA 36
DB 20 EVTQVVGVTRSRSA 36

RESULT 11
ABG65741
ID ABG65741 standard; peptide; 10 AA.
XX
AC ABG65741;
XX
DT 27-AUG-2002 (first entry)
XX
DE Plant ICK protein conserved motif 1 #49.
XX
KW Plant; inhibitor of cyclin dependent kinase; ICK.
XX
OS Zea mays.
XX
PN WO200228893-A2.
XX
PD 11-APR-2002.
XX
PF 29-JUN-2001; 2001WO-IB001492.
XX
PR 14-JUL-2000; 2000US-0218471P.
PR 13-OCT-2000; 2000US-0241219P.
XX
PA (CROP-) CROPDESIGN NV.
XX
PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
PI Hatzfeld Y;
XX
DR WPI; 2002-471311/50.
XX
PT Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
PT to screen substrates, drugs or compounds which modulate ICK activity and
PT treat disorders characterized by an insufficient or excessive production
PT of ICK inhibitors.
XX
PS Disclosure; Page 14; 141pp; English.
XX
CC This invention relates to the DNA and protein sequences of novel isolated

CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
CC the invention may be used for treating disorders characterised by
CC insufficient or excessive production of an ICK inhibitor. The protein of
CC the invention may also be used to screen for naturally-occurring ICK
CC substrates, drugs or compounds which modulate ICK activity, as well as to
CC treat disorders characterised by insufficient or excessive production of
CC ICK protein, forms which have decreased or aberrant activity compared to
CC ICK wild type protein. The present sequence represents an inhibitor of
CC cyclin dependent kinase (ICK) protein of the invention
XX
SQ Sequence 10 AA;

Query Match 3.9%; Score 10; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 212 EIQEFFFFAAAE 221
DB 1 EIQEFFFFAAAE 10

RESULT 12
ABG65692
ID ABG65692 standard; protein; 226 AA.
XX
AC ABG65692;
XX
DT 27-AUG-2002 (first entry)
XX
DE Rice OsICK5 protein.
XX
KW Plant; inhibitor of cyclin dependent kinase; ICK.
XX
OS Oryza sativa.
XX
PN WO200228893-A2.
XX
PD 11-APR-2002.
XX
PF 29-JUN-2001; 2001WO-IB001492.
XX
PR 14-JUL-2000; 2000US-0218471P.
PR 13-OCT-2000; 2000US-0241219P.
XX
PA (CROP-) CROPDESIGN NV.
XX
PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
PI Hatzfeld Y;
XX
DR WPI; 2002-471311/50.
DR N-PSDB; ABK93981.
XX
PT Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
PT to screen substrates, drugs or compounds which modulate ICK activity and
PT treat disorders characterized by an insufficient or excessive production
PT of ICK inhibitors.
XX
PS Claim 48; Page 139-140; 141pp; English.
XX
CC This invention relates to the DNA and protein sequences of novel isolated
CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
CC the invention may be used for treating disorders characterised by
CC insufficient or excessive production of an ICK inhibitor. The protein of
CC the invention may also be used to screen for naturally-occurring ICK
CC substrates, drugs or compounds which modulate ICK activity, as well as to
CC treat disorders characterised by insufficient or excessive production of
CC ICK protein, forms which have decreased or aberrant activity compared to
CC ICK wild type protein. The present sequence represents an inhibitor of
CC cyclin dependent kinase (ICK) protein of the invention
XX
SQ Sequence 226 AA;

Query Match 3.9%; Score 10; DB 5; Length 226;

Best Local Similarity 100.0%; Pred. No. 0.63; Mismatches 0; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 VRGVPLDAGG 246
Db 205 VRGVPLDAGG 214
|||||

RESULT 13
ABG65793
ID ABG65793 standard; peptide; 25 AA.
XX AC ABG65793;
XX DT 27-AUG-2002 (first entry)
XX DE Plant potential PEST sequence #9.
XX KW Plant; inhibitor of cyclin dependent kinase; ICK.
XX OS Oryza sativa.
XX PN WO200228893-A2.
XX PD 11-APR-2002.
XX PF 29-JUN-2001; 2001WO-IB001492.
XX PR 14-JUL-2000; 2000US-0218471P.
XX PR 13-OCT-2000; 2000US-0241219P.
XX PA (CROP-) CROPDESIGN NV.
XX PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
XX PI Hatzfeld Y;
XX DR WPI; 2002-471311/50.
XX PT Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
PT to screen substrates, drugs or compounds which modulate ICK activity and
PT treat disorders characterized by an insufficient or excessive production
PT of ICK inhibitors.
XX PS Disclosure; Page 17; 141pp; English.
XX CC This invention relates to the DNA and protein sequences of novel isolated
CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
CC the invention may be used for treating disorders characterised by
CC insufficient or excessive production of an ICK inhibitor. The protein of
CC the invention may also be used to screen for naturally-occurring ICK
CC substrates, drugs or compounds which modulate ICK activity, as well as to
CC treat disorders characterised by insufficient or excessive production of
CC ICK protein, forms which have decreased or aberrant activity compared to
CC ICK wild type protein. The present sequence represents an inhibitor of
CC cyclin dependent kinase (ICK) protein of the invention
XX SQ Sequence 25 AA;

Query Match 3.5%; Score 9; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 181 SDLESDLAG 189
Db 14 SDLESDLAG 22
|||||

RESULT 14
AAP01950
ID AAP01950 standard; protein; 37 AA.
XX AC AAP01950;
XX XX

DT 01-NOV-2001 (first entry)
XX Rice Cyclin dependent kinase inhibitor (CDKI) clone rds2c.pk0088.024.
DE Cyclin dependent kinase inhibitor; CDKI; herbicide; cell cycle; rice;
XX plant growth inhibitor.
KW Oryza sativa.
XX OS
XX FH Key Location/Qualifiers
FT Misc-difference 33 /note= "Encoded by GTCA"
XX FT
XX PN WO200060087-A2.
XX PD 12-OCT-2000.
XX PF 06-APR-2000; 2000WO-US009106.
XX PR 07-APR-1999; 99US-0128192P.
XX PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX PI Klein TM, Weng Z, Cahoon RE;
XX DR WPI; 2000-679375/66.
XX DR N-PSDB; AAN02400.
XX PT Cyclin dependent kinase inhibitor sequences, useful for identifying
PT herbicides and plant growth inhibitors.
XX PS Claim 10; Fig 1; 58pp; English.
XX CC The invention describes a novel isolated polynucleotide comprising a
CC nucleotide sequence encoding one of 17 specific cyclin dependent kinase
CC inhibitor (CDKI) polypeptides, cell cycle regulators involved in control
CC of cell division, growth and death. The nucleotide sequences can be used
CC in a vector to transform a host cell to produce the CDKI polypeptide.
CC They can also be used in methods for selecting and obtaining a nucleic
CC acid sequence that encodes CDKI or affects the level of CDKI expression.
CC The encoded protein can be used in a method for evaluating a compound for
CC its ability to inhibit the activity of a CDKI. The inhibitors can be used
CC as herbicides. They can also be used to inhibit plant growth. The
CC polynucleotide sequences can be used in gene mapping and as genetic
CC markers. The sequence is the rice CDKI clone rds2c.pk0088.024 as described
CC in the method of the invention
XX SQ Sequence 37 AA;

Query Match 3.5%; Score 9; DB 3; Length 37;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 VRGVPLDAG 245
Db 17 VRGVPLDAG 25
|||||

RESULT 15
AAB27252
ID AAB27252 standard; protein; 37 AA.
XX AC AAB27252;
XX DT 17-JAN-2001 (first entry)
XX DE Rice cyclin-dependent kinase inhibitor #3.
XX KW Rice; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
XX KW cell growth; herbicide.
XX OS Oryza sativa.
XX XX

FH Key Location/Qualifiers
FT Misc-difference 34 /note= "encoded by AAGC"
FT
XX
PN WO200060087-A2.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US009106.
XX
PR 07-APR-1999; 99US-0128192P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Klein TM, Weng Z, Cahoon RE;
XX
DR WPI; 2000-679375/66.
DR N-PSDB; AAA95286.
XX
PT Cyclin dependent kinase inhibitor sequences, useful for identifying
PT herbicides and plant growth inhibitors.
XX
PS Claim 10; Fig 1; 58pp; English.
XX
CC The present sequence is the rice cyclin-dependent kinase inhibitor
CC (CDKI). Its coding sequence was isolated by searching a rice seed cDNA
CC library for sequences similar to those encoding the CDKI from Chenopodium
CC rubrum, Caenorhabditis elegans and Arabidopsis thaliana. CDKI is involved
CC in the cell cycle, and may promote or inhibit cell division and growth.
CC The protein and its coding sequence are useful in the production of
CC transgenic plants which produce increased or decreased amounts of the
CC CDKI protein, in the identification of herbicides, in genetic and
CC physical mapping and in the isolation of the CDKI gene in other organisms
XX
SQ Sequence 37 AA;

Query Match 3.5%; Score 9; DB 3; Length 37;
Best Local Similarity 100.0%; Pred. NO. 1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 VRGVPLDAG 245
Db |||||
17 VRGVPLDAG 25

Search completed: October 1, 2004, 17:08:02
Job time : 128 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 17:05:59 ; Search time 18 Seconds
(without alignments)
734.236 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 256
Sequence: 1 MGKYMFKRGGAAGAEVAE.....VRGVPLDAGGRFEWAPVWSI 256

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 8 | 3.1 | 62 | 4 | US-09-295-996B-4 |
| 2 | 8 | 3.1 | 62 | 4 | US-09-295-846B-4 |
| 3 | 8 | 3.1 | 62 | 4 | US-09-551-737C-4 |
| 4 | 8 | 3.1 | 62 | 4 | US-09-551-738B-4 |
| 5 | 8 | 3.1 | 209 | 4 | US-09-685-166A-897 |
| 6 | 8 | 3.1 | 254 | 4 | US-09-252-991A-18443 |
| 7 | 8 | 3.1 | 358 | 4 | US-09-252-991A-27290 |
| 8 | 8 | 3.1 | 492 | 3 | US-09-342-749-2 |
| 9 | 8 | 3.1 | 492 | 4 | US-09-691-840-2 |
| 10 | 8 | 3.1 | 492 | 4 | US-09-685-166A-895 |
| 11 | 8 | 3.1 | 601 | 1 | US-07-676-174A-2 |
| 12 | 8 | 3.1 | 601 | 1 | US-08-194-338-2 |
| 13 | 8 | 3.1 | 671 | 4 | US-09-252-991A-23641 |
| 14 | 7 | 2.7 | 10 | 4 | US-09-755-630B-186 |
| 15 | 7 | 2.7 | 15 | 2 | US-08-592-646A-33 |
| 16 | 7 | 2.7 | 15 | 2 | US-08-592-646A-34 |
| 17 | 7 | 2.7 | 15 | 2 | US-08-592-646A-35 |
| 18 | 7 | 2.7 | 15 | 2 | US-08-592-646A-36 |
| 19 | 7 | 2.7 | 15 | 2 | US-08-592-646A-37 |
| 20 | 7 | 2.7 | 15 | 2 | US-08-592-646A-38 |
| 21 | 7 | 2.7 | 15 | 2 | US-08-592-646A-39 |
| 22 | 7 | 2.7 | 15 | 4 | US-09-165-422-33 |
| 23 | 7 | 2.7 | 15 | 4 | US-09-165-422-34 |
| 24 | 7 | 2.7 | 15 | 4 | US-09-165-422-35 |
| 25 | 7 | 2.7 | 15 | 4 | US-09-165-422-36 |
| 26 | 7 | 2.7 | 15 | 4 | US-09-165-422-37 |
| 27 | 7 | 2.7 | 15 | 4 | US-09-165-422-38 |

| | | | | | | |
|----|---|-----|-----|---|----------------------|-------------------|
| 28 | 7 | 2.7 | 15 | 4 | US-09-165-422-39 | Sequence 39, Appl |
| 29 | 7 | 2.7 | 20 | 1 | US-08-440-861-4 | Sequence 4, Appli |
| 30 | 7 | 2.7 | 33 | 1 | US-08-440-861-54 | Sequence 54, Appl |
| 31 | 7 | 2.7 | 34 | 1 | US-08-433-854-7 | Sequence 7, Appli |
| 32 | 7 | 2.7 | 34 | 1 | US-08-174-745A-7 | Sequence 7, Appli |
| 33 | 7 | 2.7 | 34 | 2 | US-08-195-947-7 | Sequence 7, Appli |
| 34 | 7 | 2.7 | 34 | 2 | US-08-433-885-7 | Sequence 7, Appli |
| 35 | 7 | 2.7 | 34 | 2 | US-08-433-908B-7 | Sequence 7, Appli |
| 36 | 7 | 2.7 | 34 | 3 | US-08-410-614-7 | Sequence 7, Appli |
| 37 | 7 | 2.7 | 37 | 2 | US-08-592-646A-59 | Sequence 7, Appli |
| 38 | 7 | 2.7 | 37 | 4 | US-07-796-361A-15 | Sequence 59, Appl |
| 39 | 7 | 2.7 | 48 | 1 | US-09-252-991A-32099 | Sequence 15, Appl |
| 40 | 7 | 2.7 | 99 | 4 | US-09-489-039A-9619 | Sequence 32099, A |
| 41 | 7 | 2.7 | 99 | 4 | US-09-489-039A-9619 | Sequence 9619, Ap |
| 42 | 7 | 2.7 | 112 | 4 | US-09-621-976-5911 | Sequence 5911, Ap |
| 43 | 7 | 2.7 | 118 | 4 | US-09-732-210-322 | Sequence 322, App |
| 44 | 7 | 2.7 | 123 | 4 | US-09-252-991A-30802 | Sequence 30802, A |
| 45 | 7 | 2.7 | 134 | 4 | US-09-252-991A-22133 | Sequence 22133, A |

ALIGNMENTS

RESULT 1
US-09-295-996B-4
; Sequence 4, Application US/09295996B
; Patent No. 6413530
; GENERAL INFORMATION:
; APPLICANT: Borovsky, Dov
; TITLE OF INVENTION: PESTICIDAL PEPTIDES
; FILE REFERENCE: UF-230
; CURRENT APPLICATION NUMBER: US/09/295,996B
; CURRENT FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Aedes aegypti
US-09-295-996B-4

Query Match 3.1%; Score 8; DB 4; Length 62;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 SAASNSGS 161
Db 38 SAASNSGS 45

RESULT 2
US-09-295-846B-4
; Sequence 4, Application US/09295846B
; Patent No. 6562590
; GENERAL INFORMATION:
; APPLICANT: Borovsky, Dov
; TITLE OF INVENTION: Transformed Cells Useful for the Control of Pests
; FILE REFERENCE: UF-223
; CURRENT APPLICATION NUMBER: US/09/295,846B
; CURRENT FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Aedes aegypti
US-09-295-846B-4

Query Match 3.1%; Score 8; DB 4; Length 62;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 SAASNSGS 161

Db 38 SAASNSGS 45
|||||

RESULT 3
US-09-551-737C-4
; Sequence 4, Application US/09551737C
; Patent No. 6566129
; GENERAL INFORMATION:
; APPLICANT: Borovsky, Dov
; APPLICANT: Schlesinger, Yaagov
; APPLICANT: Nauwelaers, Sabine M. I.
; TITLE OF INVENTION: Transformed Cells Useful for the Control of Pests
; FILE REFERENCE: UF-223C1
; CURRENT APPLICATION NUMBER: US/09/551,737C
; CURRENT FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/295,846
; PRIOR FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Aedes aegypti
US-09-551-737C-4

Query Match 3.1%; Score 8; DB 4; Length 62;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 SAASNSGS 161
|||||
Db 38 SAASNSGS 45

RESULT 4
US-09-551-738B-4
; Sequence 4, Application US/09551738B
; Patent No. 6635265
; GENERAL INFORMATION:
; APPLICANT: Borovsky, Dov
; TITLE OF INVENTION: Materials and Methods Useful for the Control of Insect Larvae
; FILE REFERENCE: UF-224C1
; CURRENT APPLICATION NUMBER: US/09/551,738B
; CURRENT FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/296,113
; PRIOR FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Aedes aegypti
US-09-551-738B-4

Query Match 3.1%; Score 8; DB 4; Length 62;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 SAASNSGS 161
|||||
Db 38 SAASNSGS 45

RESULT 5
US-09-685-166A-897
; Sequence 897, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 897
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-685-166A-897

Query Match 3.1%; Score 8; DB 4; Length 209;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 GAALAAAGL 113
|||||
Db 97 GAALAAAGL 104

RESULT 6
US-09-252-991A-18443
; Sequence 18443, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18443
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18443

Query Match 3.1%; Score 8; DB 4; Length 254;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 GAAGQOQA 107
|||||
Db 199 GAAGQOQA 206

RESULT 7
US-09-252-991A-27290
; Sequence 27290, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27290
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27290

Query Match 3.1%; Score 8; DB 4; Length 358;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 GAALAAAGL 113
Db 334 GAALAAAGL 341

RESULT 8

US-09-342-749-2
; Sequence 2, Application US/09342749
; Patent No. 6166194

; GENERAL INFORMATION:

; APPLICANT: Wong, Alexander K.C.
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.-F.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: TMRSS2 is a Tumor Suppressor
; FILE REFERENCE: 2318-202
; CURRENT APPLICATION NUMBER: US/09/342,749
; CURRENT FILING DATE: 1999-06-29
; EARLIER APPLICATION NUMBER: US 60/091,044
; EARLIER FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2

; LENGTH: 492
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-342-749-2

Query Match 3.1%; Score 8; DB 3; Length 492;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 GAALAAAGL 113
Db 97 GAALAAAGL 104

RESULT 9

US-09-691-840-2
; Sequence 2, Application US/09691840
; Patent No. 644419

; GENERAL INFORMATION:

; APPLICANT: Wong, Alexander K.C.
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.-F.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: TMRSS2 is a Tumor Suppressor
; FILE REFERENCE: 2318-202
; CURRENT APPLICATION NUMBER: US/09/691,840
; CURRENT FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: US/09/342,749
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 60/091,044
; PRIOR FILING DATE: 1998-06-29

; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-691-840-2

Query Match 3.1%; Score 8; DB 4; Length 492;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 GAALAAAGL 113
Db 97 GAALAAAGL 104

RESULT 10

US-09-685-166A-895
; Sequence 895, Application US/09685166A
; Patent No. 6630305

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 895

; LENGTH: 492
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-685-166A-895

Query Match 3.1%; Score 8; DB 4; Length 492;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 GAALAAAGL 113
Db 97 GAALAAAGL 104

RESULT 11

US-07-676-174A-2
; Sequence 2, Application US/07676174A
; Patent No. 5344776

; GENERAL INFORMATION:

; APPLICANT: Vantor, J. Craig
; TITLE OF INVENTION: Octopamine receptor
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: Eleventh Floor, 1615 L. Street, N.W.
; CITY: Washington

STATE: D.C.
COUNTRY: USA
ZIP: 20036-5601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/676,174A
FILING DATE: 19910328
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Scott, Watson T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: WTS/5683/79117/KIK
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)861-3000
TELEFAX: (202)822-8944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 601 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-676-174A-2

Query Match 3.1%; Score 8; DB 1; Length 601;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 AGDGGSC 68
| | | | |
Db 435 AGDGGSC 442

RESULT 12
US-08-194-338-2
Sequence 2, Application US/08194338
Patent No. 5474898
GENERAL INFORMATION:
APPLICANT: Venter, John C.
APPLICANT: Fraser, Claire M.
APPLICANT: McCombie, William R.
TITLE OF INVENTION: OCTOPAMINE RECEPTOR
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson and Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/194,338
FILING DATE: 08-FEB-1994
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: US 07/676,174
FILING DATE: 28-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH101.001DV1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 601 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-194-338-2

Query Match 3.1%; Score 8; DB 1; Length 601;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 AGDGGSC 68
| | | | |
Db 435 AGDGGSC 442

RESULT 13
US-09-252-991A-23641
Sequence 23641, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 23641
LENGTH: 671
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23641

Query Match 3.1%; Score 8; DB 4; Length 671;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 ADGAAGQ 105
| | | | |
Db 522 ADGAAGQ 529

RESULT 14
US-09-755-630B-186
Sequence 186, Application US/09755630B
Patent No. 6639054
GENERAL INFORMATION:
APPLICANT: ALIBHAI, MURTAZA F.
APPLICANT: ASTWOOD, JAMES D.
APPLICANT: SAMPSON, HUGH A.
APPLICANT: MCWHERTER, CHARLES A.
TITLE OF INVENTION: PREPARATION OF DEALLERGENIZED PROTEINS AND PERMUTEINS
FILE REFERENCE: 11899.0217.NFUS00 (MOBT217)
CURRENT APPLICATION NUMBER: US/09/755,630B
CURRENT FILING DATE: 2001-01-05
PRIOR APPLICATION NUMBER: US 60/174,669
PRIOR FILING DATE: 2000-01-06
NUMBER OF SEQ ID NOS: 295
SOFTWARE: PatentIn version 3.0
SEQ ID NO 186
LENGTH: 10
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic polypeptide
US-09-755-630B-186

Query Match 2.7%; Score 7; DB 4; Length 10;
 Best Local Similarity 100.0%; Pred. No. 7;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 FAAAAA 223
 Db 3 FAAAAA 9

RESULT 15
 US-08-592-646A-33
 ; Sequence 33, Application US/08592646A
 ; Patent No. 5851535
 ; GENERAL INFORMATION:
 ; APPLICANT: JOLIVET-REYNAUD, Colette
 ; TITLE OF INVENTION: MIMOTOPIC POLYPEPTIDES OF TOXOPLASMA
 ; TITLE OF INVENTION: GONDII AND APPLICATIONS
 ; NUMBER OF SEQUENCES: 63
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: OLIFF & BERRIDGE, PLC
 ; STREET: P.O. Box 19928
 ; CITY: Alexandria
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22320
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/592,646A
 ; FILING DATE: 26-JAN-1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Berridge, William P.
 ; REGISTRATION NUMBER: 30,024
 ; REFERENCE/DOCKET NUMBER: WPB 36923
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 703 836-6400
 ; TELEFAX: 703 836-2787
 ; INFORMATION FOR SEQ ID NO: 33:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 15 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: not relevant
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ;
 US-08-592-646A-33

Query Match 2.7%; Score 7; DB 2; Length 15;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 RGAAGAE 15
 Db 9 RGAAGAE 15

Search completed: October 1, 2004, 17:21:57
 Job time : 19 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 17:10:40 ; Search time 131 Seconds
(without alignments)
628.859 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 256
Sequence: 1 MGKYMRCRGAAGAEVAE.....VRGVLDAAGRFEWAPVVISI 256

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1351062 seqs, 321799191 residues

Word size : 0
Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 256 | 100.0 | 256 | 9 | US-09-993-308-2 |
| 2 | 256 | 100.0 | 256 | 11 | US-09-993-808B-2 |
| 3 | 117 | 45.7 | 221 | 12 | US-10-425-114-69245 |
| 4 | 117 | 45.7 | 263 | 12 | US-10-425-114-63204 |
| 5 | 21 | 8.2 | 116 | 15 | US-10-333-006-14 |
| 6 | 17 | 6.6 | 227 | 9 | US-09-993-308-6 |
| 7 | 17 | 6.6 | 227 | 11 | US-09-993-808B-6 |
| 8 | 17 | 6.6 | 262 | 15 | US-10-333-006-10 |
| 9 | 17 | 6.6 | 278 | 12 | US-10-425-114-71211 |
| 10 | 17 | 6.6 | 417 | 16 | US-10-437-963-195115 |
| 11 | 10 | 3.9 | 226 | 15 | US-10-333-006-55 |
| 12 | 10 | 3.9 | 354 | 16 | US-10-437-963-122432 |
| 13 | 10 | 3.9 | 423 | 16 | US-10-437-963-122443 |
| 14 | 9 | 3.5 | 324 | 12 | US-10-107-431-13 |
| 15 | 8 | 3.1 | 62 | 13 | US-10-090-696-2 |

| | | | | | |
|----|-----|-----|----|----------------------|--------------------|
| 16 | 3.1 | 150 | 16 | US-10-437-963-112674 | Sequence 112674, A |
| 17 | 3.1 | 155 | 12 | US-10-425-114-42206 | Sequence 42206, A |
| 18 | 3.1 | 165 | 16 | US-10-437-963-154569 | Sequence 154569, A |
| 19 | 3.1 | 173 | 16 | US-10-437-963-200561 | Sequence 200561, A |
| 20 | 3.1 | 177 | 16 | US-10-437-963-139291 | Sequence 139291, A |
| 21 | 3.1 | 205 | 16 | US-10-437-963-140830 | Sequence 140830, A |
| 22 | 3.1 | 209 | 9 | US-09-759-143-897 | Sequence 897, App |
| 23 | 3.1 | 209 | 9 | US-09-780-669-897 | Sequence 897, App |
| 24 | 3.1 | 209 | 9 | US-09-822-827-897 | Sequence 897, App |
| 25 | 3.1 | 209 | 9 | US-09-895-793-897 | Sequence 897, App |
| 26 | 3.1 | 209 | 9 | US-09-895-814-897 | Sequence 897, App |
| 27 | 3.1 | 209 | 13 | US-10-012-896-897 | Sequence 897, App |
| 28 | 3.1 | 209 | 14 | US-10-144-678A-897 | Sequence 897, App |
| 29 | 3.1 | 209 | 14 | US-10-294-025-897 | Sequence 897, App |
| 30 | 3.1 | 221 | 15 | US-10-369-493-12041 | Sequence 12041, A |
| 31 | 3.1 | 234 | 15 | US-10-421-138A-128 | Sequence 128, App |
| 32 | 3.1 | 234 | 15 | US-10-374-780A-1242 | Sequence 1242, App |
| 33 | 3.1 | 316 | 15 | US-10-084-846A-116 | Sequence 116, App |
| 34 | 3.1 | 336 | 12 | US-10-282-122A-55536 | Sequence 55536, A |
| 35 | 3.1 | 427 | 12 | US-10-282-122A-60524 | Sequence 60524, A |
| 36 | 3.1 | 491 | 14 | US-10-156-761-13382 | Sequence 13382, A |
| 37 | 3.1 | 492 | 9 | US-09-759-143-895 | Sequence 895, App |
| 38 | 3.1 | 492 | 9 | US-09-759-143-932 | Sequence 932, App |
| 39 | 3.1 | 492 | 9 | US-09-780-669-895 | Sequence 895, App |
| 40 | 3.1 | 492 | 9 | US-09-780-669-932 | Sequence 932, App |
| 41 | 3.1 | 492 | 9 | US-09-879-792-14 | Sequence 14, Appl |
| 42 | 3.1 | 492 | 9 | US-09-822-827-895 | Sequence 895, App |
| 43 | 3.1 | 492 | 9 | US-09-822-827-932 | Sequence 932, App |
| 44 | 3.1 | 492 | 9 | US-09-895-793-895 | Sequence 895, App |
| 45 | 3.1 | 492 | 9 | US-09-895-793-932 | Sequence 932, App |

ALIGNMENTS

RESULT 1
US-09-993-308-2
; Sequence 2, Application US/09993308
; Patent No. US20020159435A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William J.
; APPLICANT: Lowe, Keith S.
; APPLICANT: Larkins, Brian A.
; APPLICANT: Dilkes, Brian R.
; APPLICANT: Sun, Yuejin
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993.308
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 60/246,349
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Zea mays
US-09-993-308-2

| | | | |
|----|-----|---|-----|
| QY | 1 | MGKYMRCRGAAGAEVAEVTQVVGVRTRSRSAATGGVAKVAPRRKRAPAGEPAAAVS | 60 |
| Db | 1 | MGKYMRCRGAAGAEVAEVTQVVGVRTRSRSAATGGVAKVAPRRKRAPAGEPAAAVS | 60 |
| QY | 61 | AGDGGSCYIHLRSRMLFMAPPQPSVDSVPTPVEAADGAAGQGAALAAGLSRCSSSTA | 120 |
| Db | 61 | AGDGGSCYIHLRSRMLFMAPPQPSVDSVPTPVEAADGAAGQGAALAAGLSRCSSSTA | 120 |
| QY | 121 | SSVNLGLGGQSGSHTCRSYDAAEAGGDHVLVDVSAASNSGSPDRRETTPSSRAHGEL | 180 |

Db 121 SSVNGLGGQGSHTCRSYDAAEAGGDHVLVDVSAASNSGSGPDRERRETTTPSSRAHGEL 180
QY 181 SDLESDLAGHKTGPSLPAATPAAEELIVPPAHEIQEFFFFAAAEAAQAQKRFASKYNFDFVRGV 240
Db 181 SDLESDLAGHKTGPSLPAATPAAEELIVPPAHEIQEFFFFAAAEAAQAQKRFASKYNFDFVRGV 240
QY 241 PLDAGGRFEWAPVVS 256
Db 241 PLDAGGRFEWAPVVS 256

RESULT 2
US-09-993-808B-2
; Sequence 2, Application US/09993808B
; Publication No. US20040003433A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William
; APPLICANT: Lowe, Keith
; APPLICANT: Sun, Yuejin
; APPLICANT: Dilkes, Brian
; APPLICANT: Larkins, Brian
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides,
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,808B
; CURRENT FILING DATE: 2001-11-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 256
; TYPE: PRT
; ORGANISM: zea mays
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(256)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-993-808B-2

Query Match 100.0%; Score 256; DB 11; Length 256;
Best Local Similarity 100.0%; Pred. No. 3.le-224;
Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGKYMRCRGAGAEVAAVEVTQVGVTRSRSAATGGVAKVAPRRKRAPAGEPAAAVS 60
Db 1 MGKYMRCRGAGAEVAAVEVTQVGVTRSRSAATGGVAKVAPRRKRAPAGEPAAAVS 60
QY 61 AGDGGSCYIHLRSRMFLMAPPQPPSVDSVPTPVEAADGAAGQQAALAGLSRCSSTA 120
Db 61 AGDGGSCYIHLRSRMFLMAPPQPPSVDSVPTPVEAADGAAGQQAALAGLSRCSSTA 120
QY 121 SSVNGLGGQGSHTCRSYDAAEAGGDHVLVDVSAASNSGSGPDRERRETTTPSSRAHGEL 180
Db 121 SSVNGLGGQGSHTCRSYDAAEAGGDHVLVDVSAASNSGSGPDRERRETTTPSSRAHGEL 180
QY 181 SDLESDLAGHKTGPSLPAATPAAEELIVPPAHEIQEFFFFAAAEAAQAQKRFASKYNFDFVRGV 240
Db 181 SDLESDLAGHKTGPSLPAATPAAEELIVPPAHEIQEFFFFAAAEAAQAQKRFASKYNFDFVRGV 240
QY 241 PLDAGGRFEWAPVVS 256
Db 241 PLDAGGRFEWAPVVS 256

RESULT 3
US-10-425-114-69245
; Sequence 69245, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 69245
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3079-036-F9_FLI.pep
US-10-425-114-69245

Query Match 45.7%; Score 117; DB 12; Length 221;
Best Local Similarity 100.0%; Pred. No. 6e-98;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 140 DAAEAGGDHVLVDVSAASNSGSGPDRERRETTTPSSRAHGELESDLAGHKTGPSLPAA 199
Db 105 DAAEAGGDHVLVDVSAASNSGSGPDRERRETTTPSSRAHGELESDLAGHKTGPSLPAA 164
QY 200 TPAEELIVPPAHEIQEFFFFAAAEAAQAQKRFASKYNFDFVRGVPDAGGRFEWAPVVS 256
Db 165 TPAEELIVPPAHEIQEFFFFAAAEAAQAQKRFASKYNFDFVRGVPDAGGRFEWAPVVS 221

RESULT 4
US-10-425-114-63204
; Sequence 63204, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 63204
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17355F05_FLI.pep
US-10-425-114-63204

Query Match 45.7%; Score 117; DB 12; Length 263;
Best Local Similarity 100.0%; Pred. No. 7e-98;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 140 DAAEAGGDHVLVDVSAASNSGSGPDRERRETTTPSSRAHGELESDLAGHKTGPSLPAA 199
Db 147 DAAEAGGDHVLVDVSAASNSGSGPDRERRETTTPSSRAHGELESDLAGHKTGPSLPAA 206
QY 200 TPAEELIVPPAHEIQEFFFFAAAEAAQAQKRFASKYNFDFVRGVPDAGGRFEWAPVVS 256
Db 207 TPAEELIVPPAHEIQEFFFFAAAEAAQAQKRFASKYNFDFVRGVPDAGGRFEWAPVVS 263

RESULT 5
US-10-333-006-14
; Sequence 14, Application US/10333006
; Publication No. US20040019926A1
; GENERAL INFORMATION:
; APPLICANT: Frankard, Valerie Marie-No. US20040019926A1lle S.
; APPLICANT: Peres Bota, Adrian Marius

RESULT, T 7

```
RESULT 9
US-10-425-114-71211
; Sequence 71211, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 71211
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3078-022-G5_FLI.pep
US-10-425-114-71211

Query Match          6.6%; Score 17; DB 12; Length 278;
Best Local Similarity 100.0%; Pred. No. 5.8e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      23 QVGVTRRSRSAAATGG 39
      |||||
Db      69 QVGVTRRSRSAAATGG 85

RESULT 10
US-10-437-963-195115
; Sequence 195115, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 195115
; LENGTH: 417
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91095C.1.pep
US-10-437-963-195115

Query Match          6.6%; Score 17; DB 16; Length 417;
Best Local Similarity 100.0%; Pred. No. 8.2e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      20 ETVGVGVTRRSRSAAA 36
      |||||
Db      175 ETVGVGVTRRSRSAAA 191

RESULT 11
US-10-333-006-55
; Sequence 55, Application US/10333006
```

```
; Publication No. US20040019926A1
; GENERAL INFORMATION:
; APPLICANT: Frankard, Valerie Marie-No. US20040019926A111e S.
; APPLICANT: Peres Bota, Adrian Marius
; APPLICANT: Droual, Anne-Marie
; APPLICANT: Mironov, Vladimir
; APPLICANT: Inz, Dirk
; APPLICANT: Hatzfeld, Yves
; TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
; FILE REFERENCE: 1187-13
; CURRENT APPLICATION NUMBER: US/10/333,006
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: PCT/IB01/01492
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/218,471
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/241,219
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: probe or primer
; NAME/KEY: misc_feature
; LOCATION: (119)..(119)
; OTHER INFORMATION: Xaa = any amino acid
US-10-333-006-55

Query Match          3.9%; Score 10; DB 15; Length 226;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      237 VRGVPLDAGG 246
      |||||
Db      205 VRGVPLDAGG 214

RESULT 12
US-10-437-963-122432
; Sequence 122432, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 122432
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_2535C.1.pep
US-10-437-963-122432

Query Match          3.9%; Score 10; DB 16; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      237 VRGVPLDAGG 246
```

Db 228 VRGVPLDAGG 237
|||||

RESULT 13

US-10-437-963-122443
; Sequence 122443, Application US/10437963
; Publication No. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 122443

; LENGTH: 423

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_2536C.1.pep
US-10-437-963-122443

Query Match 3.9%; Score 10; DB 16; Length 423;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 VRGVPLDAGG 246
|||||

Db 241 VRGVPLDAGG 250

RESULT 14

US-10-107-431-13

; Sequence 13, Application US/10107431

; Publication No. US20030224364A1

; GENERAL INFORMATION:

; APPLICANT: Farnet, Chris

; APPLICANT: Staffa, Alfredo

; APPLICANT: Zazopoulos, Emmanuel

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING AND DISTINGUISHING ORTHO

; TITLE OF INVENTION: BIOSYNTHETIC LOCI

; FILE REFERENCE: 3001-7US

; CURRENT APPLICATION NUMBER: US/10/107,431

; CURRENT FILING DATE: 2002-03-28

; NUMBER OF SEQ ID NOS: 282

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 13

; LENGTH: 324

; TYPE: PRT

; ORGANISM: Streptomyces mobaraensis

US-10-107-431-13

Query Match 3.5%; Score 9; DB 12; Length 324;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 AAGAEVAAV 19
|||||

Db 27 AAGAEVAAV 35

RESULT 15

US-10-090-696-2

; Sequence 2, Application US/10090696

; Publication No. US20020132302A1
; GENERAL INFORMATION:
; APPLICANT: Borovsky, Dov
; TITLE OF INVENTION: TMOF RECEPTOR AND USES THEREOF
; FILE REFERENCE: UF-214P
; CURRENT APPLICATION NUMBER: US/10/090,696
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: US/09/201,568
; PRIOR FILING DATE: 1998-11-30
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Aedes aegypti
US-10-090-696-2

Query Match 3.1%; Score 8; DB 13; Length 62;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 SAASNSGS 161
|||||

Db 38 SAASNSGS 45

Search completed: October 1, 2004, 17:32:37
Job time : 133 secs

This Page Blank (usr)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2004, 09:47:45 ; Search time 2471 Seconds
(without alignments)
3093.775 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 1304
Sequence: 1 MGKYMRCRGAGAEVAAVE.....VRGVPLDAGGRFEWAPVVISI 256

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US0993808/runat_01102004_171702_13035/app_query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0993808@cgn_1_1_5180@runat_01102004_171702_13035 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|-------|-------|--------|----|----------|--------------------|
| 1 | 820.5 | 62.9 | 1197 | 11 | AY108168 | AY108168 Zea mays |
| 2 | 751.5 | 57.6 | 659 | 13 | CA074449 | CA074449 SCEZAM108 |
| C 3 | 744.5 | 57.1 | 655 | 13 | CA104638 | CA104638 SCJFHR103 |
| 4 | 690.5 | 53.0 | 705 | 13 | CA106521 | CA106521 SCQGHR101 |
| 5 | 660 | 50.6 | 661 | 14 | CA256119 | CA256119 SCJLFL418 |
| C 6 | 656 | 50.3 | 666 | 29 | CC610974 | CC610974 OGVFR67TV |
| C 7 | 655.5 | 50.3 | 869 | 28 | BZ743686 | BZ743686 OGFAC93TC |
| 8 | 635 | 48.7 | 526 | 14 | CD444277 | CD444277 EL01N0438 |
| C 9 | 590.5 | 45.3 | 838 | 28 | CC440801 | CC440801 PUHLN53TB |
| 10 | 582.5 | 44.7 | 645 | 29 | CG136784 | CG136784 PUIDC25TB |
| 11 | 563.5 | 43.2 | 753 | 14 | CB646893 | CB646893 OSJNEB09J |
| 12 | 563 | 43.2 | 707 | 11 | AY110319 | AY110319 Zea mays |
| 13 | 522 | 40.0 | 585 | 13 | CA104568 | CA104568 SCJFHR103 |
| C 14 | 518 | 39.7 | 642 | 13 | CA104689 | CA104689 SCJFHR103 |
| 15 | 517.5 | 39.7 | 644 | 29 | CC729680 | CC729680 OGUBY35TV |
| 16 | 517 | 39.6 | 986 | 28 | CC002416 | CC002416 PUGKB53TB |
| 17 | 514 | 39.4 | 584 | 13 | CA108352 | CA108352 SCSFHR104 |
| 18 | 512 | 39.3 | 510 | 14 | CB929335 | CB929335 ABA1_41_H |
| 19 | 502 | 38.5 | 563 | 13 | CA108367 | CA108367 SCSFHR104 |
| C 20 | 498.5 | 38.2 | 987 | 28 | CC002419 | CC002419 PUGKB53TD |
| 21 | 489 | 37.5 | 890 | 29 | CG218373 | CG218373 OG1DJ10TV |
| 22 | 483 | 37.0 | 1010 | 28 | CC390569 | CC390569 PUHBT76TB |
| 23 | 474.5 | 36.4 | 531 | 14 | CA297189 | CA297189 SCUTSD208 |
| 24 | 458 | 35.1 | 697 | 14 | CB646543 | CB646543 OSJNEB09B |
| 25 | 447 | 34.3 | 823 | 29 | CC703195 | CC703195 OGVB56TH |
| 26 | 439.5 | 33.7 | 791 | 29 | CG324885 | CG324885 OGXDNO4TV |
| 27 | 439.5 | 33.7 | 917 | 29 | CG265481 | CG265481 OG2CM60TV |
| C 28 | 435 | 33.4 | 708 | 28 | BZ404098 | BZ404098 OGACI23TM |
| C 29 | 435 | 33.4 | 795 | 29 | CG339703 | CG339703 OG1DH12TH |
| 30 | 427 | 32.7 | 787 | 28 | BZ404090 | BZ404090 OGACI23TC |
| C 31 | 399.5 | 30.6 | 568 | 9 | AI737717 | AI737717 605040C07 |
| C 32 | 391.5 | 30.0 | 753 | 29 | CC729673 | CC729673 OGUBY35TH |
| C 33 | 389.5 | 29.9 | 547 | 28 | BH777311 | BH777311 fzm013f0 |
| C 34 | 389 | 29.8 | 550 | 12 | BG267004 | BG267004 1000110B0 |
| C 35 | 389 | 29.8 | 825 | 29 | CG265472 | CG265472 OG2CM60TH |
| 36 | 375.5 | 28.8 | 707 | 29 | CG358355 | CG358355 OGWIU21TH |
| C 37 | 371.5 | 28.5 | 714 | 28 | BZ642267 | BZ642267 OGAOU33TM |
| 38 | 336.5 | 25.8 | 456 | 9 | AU172982 | AU172982 AU172982 |
| 39 | 307 | 23.5 | 467 | 13 | BQ578371 | BQ578371 WHE0302_B |
| C 40 | 306.5 | 23.5 | 742 | 14 | CB646894 | CB646894 OSJNEB09J |
| C 41 | 298.5 | 22.9 | 727 | 14 | CB646544 | CB646544 OSJNEB09B |
| 42 | 290.5 | 22.3 | 642 | 28 | BZ642255 | BZ642255 OGAOU33TC |
| C 43 | 286 | 21.9 | 837 | 29 | CG030519 | CG030519 PUFTA23TB |
| 44 | 286 | 21.9 | 904 | 29 | CG030521 | CG030521 PUFTA23TD |
| C 45 | 273.5 | 21.0 | 959 | 29 | CC703206 | CC703206 OGVB56TV |

ALIGNMENTS

RESULT 1
AY108168
LOCUS AY108168 1197 bp mRNA linear HTC 16-OCT-2002
DEFINITION Zea mays PC0119368 mRNA sequence.
ACCESSION AY108168
VERSION AY108168.1 GI:21211246
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 1197)

AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
JOURNAL Overgo Probes
REFERENCE Unpublished (2002)
AUTHORS 2 (bases 1 to 1197)
TITLE Coe,E.H.
JOURNAL Direct Submission
COMMENT Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.

FEATURES
source
1. .1197
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:636945"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN

Alignment Scores:
Pred. No.: 1.09e-43 Length: 1197
Score: 820.50 Matches: 204
Percent Similarity: 77.01% Conservative: 7
Best Local Similarity: 74.45% Mismatches: 34
Query Match: 62.92% Indels: 29
DB: 11 Gaps: 17

US-09-993-808B-2 (1-256) x AY108168 (1-1197)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
155 ATGGGAAGTACATGCGAAGCGCAGGGGGCGCGGAGGGGTGGCGGAGTGGCGAGTGGAG 214
QY 21 ValThrGlnValGlyValArgThrArgSerArgSerAlaAlaThr---GlyGly 39
215 GTCTCGAGGTGCTCGGCTCCGACGAGGTCAGGTCAGGTCGCGCGGCGGCGAGCCGCGCGGT 274
QY 40 ValAlaLysValAla---ProArgArgLysArgAla-----ProAla----- 52
275 GTCGGAAGTCTGCTCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 334
QY 53 -----GlyGluProAlaAlaAlaValSerAlaGlyGlyAspGlyGlySer---CysTyr 69
335 ACGTCGGGGAGCCTGGTGGCGCTGGTGGTGGGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAG 394
QY 70 IleHisLeuArgSerArgMetLeuPheMetAlaProGlnProGlnProSerValAsp 89
395 ATCCACCTGCGGAGCGCATGCTGTTTCATGGCAGCACCTCAGCAGCAACCGTCG----- 448
QY 90 SerValProThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGly---AlaAla 108
449 GCGGCTGTGACGCGGCTGGAGGTGCT---GGTGGGCGACAGCAGGGCGGGGTGGTGGCG 505
QY 109 LeuAlaAlaGlyLeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGly 128
506 CTGCGGCTGGCCCTCTCGCGTGTGCTCCAGCAGCGGCTCGTGGTGGAGCTC----- 556
QY 129 GlyGlnArgGlySerHisThrCysArgSerTyrAspAla---AlaGluAlaGlyGlyAsp 147

Db 557 -----GGGGGCCACGCCTGCCGCTCCGACGCTGCCGCTGCCGAGGTTGACGGGGAT 607
QY 148 HisValLeuValAspValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArg 167
Db 608 CACGTC---CCGGATGTCGTACCCGAGCAACTCGGGGAGCGTCCCGGACCGGAGAGG 664
QY 168 ArgGluThrThrPro---SerSerArgAlaHis---GlyGluLeuSerAspLeuGluSer 185
Db 665 AGAGAGACGACGCCCATCGTCGAGCCGGCGGCGACGCGGCGGAGCTCAGCATCTGGAGTCG 724
QY 186 AspLeuAlaGly---HisLysThrGly---ProSerLeuProAlaAlaThrProAla--- 202
Db 725 GATCTGGTGGGCGGCGAGAGACTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 784
QY 203 AlaGluLeuLeuValProProAlaHisGluLeuGlnGluPhePheAlaAlaAlaGluAla 222
Db 785 GCGGAGCTGATGTCGCGCGGCGACAGGAGATCCAGGAATTCTTCGCGGCGCGGAGGCG 844
QY 223 AlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyValProLeu 242
Db 845 GCCCATGCCCAACGCTTGTCTTCAAGTACAACTTCGACTTCGTCGCGGCGGCTGCCCTC 904
QY 243 AspAlaGlyGlyArgPheGluTrpAlaProValValSerIle 256
Db 905 GACGCC---GGCCGTTTCGAGTGGACGCCGCGGCGGTCAGCATC 943

RESULT 2

CA074449

LOCUS

DEFINITION

5', mRNA sequence.

ACCESSION

CA074449

VERSION

CA074449.1

KEYWORDS

EST.

SOURCE

Saccharum officinarum

ORGANISM

Saccharum officinarum

REFERENCE

1 (bases 1 to 659)

AUTHORS

Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.

TITLE

The libraries that made SUCEST

JOURNAL

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

COMMENT

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found

through the Brazilian Clone Collection Center (BCCC) at

http://www.bcccenter.fcav.unesp.br

Plate: 081 row: D column: 06

Seq primer: T7 Promoter Primer.

FEATURES

Location/Qualifiers

1..659

/organism="Saccharum officinarum"

/mol_type="mRNA"

/db_xref="taxon:4547"

/clone="SCEZAM1081D06"

/lab_host="DH10B"

/clone_lib="AM1"

/note="Organ: Apical meristem and tissues surrounding of

mature plants; Vector: pSport1; Site_1: Sali; Site_2:

NotI; An unidirectional cDNA library generated from

[Apical meristem and tissues surrounding of mature

plants]. cDNA was prepared from polyA+ mRNA using

SuperScript plasmid System Kit (Invitrogen). The

double-strand cDNAs were fractionated in a sepharose

CL-2B 40cm-columns and fragments sizing between 0.8 and

1.5 Kb were directionally cloned into the vector. Details

of each source of RNA and library construction can be

obtained at <http://sucest.lad.ic.unicamp.br/public>

ORIGIN

Alignment Scores:

Pred. No.: 1.61e-39 Length: 659
Score: 751.50 Matches: 173
Percent Similarity: 84.21% Conservative: 3
Best Local Similarity: 82.78% Mismatches: 14
Query Match: 57.63% Indels: 19
DB: 13 Gaps: 11

US-09-993-808B-2 (1-256) x CA074449 (1-659)

```
QY 1 MetGlyLysTyrMetArg---LysCysArgGlyAlaAlaGlyAlaGluValAlaVal 19
Db 59 ATGGGAAGTACATCGCAGCAAGTGCAGGGGCGCGCGGAGGAGTCCGCCCGCTC 118
QY 20 GluValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaThr---Gly 38
Db 119 GAGGTACGCAGGTCTCGCGCTCCGACGAGGTCCAGGTCCGCGCGGCGACCGGGGT 178
QY 39 GlyValAlaLysValAlaProArgArgLysArgAla-----ProAla 52
Db 179 GGTGTGGGAAGTTCGACCCAGGAGGAAGAGCGCTGGCGCGCGGCCAACGTGGCGCG 238
QY 53 GlyGluProAlaAlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeu 72
Db 239 GGGGAGCCT---GCTGCCGTCCGCGCTGGAGGGGACGGCGGAAAGTGTCTACATCTG 295
QY 73 ArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAspSerValPro 92
Db 296 CGCAGCCGATGCTGTTTCATGGCACCGCCTCAGCCGCGAGCGG-----TCTGTTCCG 346
QY 93 ThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaGly 112
Db 347 ACGCCGCGGAGGCTGCTGTGGA-----GGACAGCAGGCGCGCGCTCGCGCGCGG 400
QY 113 LeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyGlyGlnArgGly 132
Db 401 CTCTCGCTNGTCCAGCACGGCGTCTGCTGGTGGACGTCCGG---GGTCATCAGCACAGG 457
QY 133 SerHisThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuValAsp 152
Db 458 AGCCAGCCTGCCGCTCC---GACGCTGCGGAGTTGGCGGNGATCACGCC---CCGAT 511
QY 153 ValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgArgGluThrPro 172
Db 512 GTCTCG---GCGAGCAACTCGGGAGCGGCCAGACCGCGAGAGAGAGACGACGCA 568
QY 173 SerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHisLys--- 191
Db 569 TCGAGCCGGCGCACGGCGAGTCTCAGCATCTGAGTCCGATCTGGCGGCGGCAACAAGAT 628
QY 192 ThrGlyProSerLeuProAlaAlaThr 200
Db 629 ACTGGCGGTCTGTTGCCGCGGCAACG 655
```

RESULT 3

CA104638/c

LOCUS

DEFINITION

SCJFHR1034B05.b HRI1 Saccharum officinarum cDNA clone SCJFHR1034B05

3', mRNA sequence.

ACCESSION

CA104638

VERSION

CA104638.1 GI:34957945

KEYWORDS

SOURCE

ORGANISM

Saccharum officinarum

Saccharum officinarum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Saccharum.

1 (bases 1 to 655)

REFERENCE

AUTHORS

Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.

TITLE

The libraries that made SUCEST

JOURNAL

COMMENT

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
<http://www.bcccenter.fcav.unesp.br>
Plate: 034 row: B column: 05
Seq primer: SP6 Promoter primer.
Location/Qualifiers
1. .655
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCJFHR1034B05"
/lab_host="DH10B"
/clone_lib="HRI1"
/note="Organ: seedlings inoculated with Herbaspirillum
rubrisubalbicans; Vector: pSport1; Site 1: SalI; Site 2:
NotI; An unidirectional cDNA library generated from
[seedlings inoculated with Herbaspirillum
rubrisubalbicans]. cDNA was prepared from polyA+ mRNA
using Superscript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at <http://sucest.lad.ic.unicamp.br/public>

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.: 4.57e-39 Length: 655
Score: 744.50 Matches: 172
Percent Similarity: 80.09% Conservative: 9
Best Local Similarity: 76.11% Mismatches: 26
Query Match: 57.09% Indels: 19
DB: 13 Gaps: 11

US-09-993-808B-2 (1-256) x CA104638 (1-655)

```
QY 1 MetGlyLysTyrMet---ArgLysCysArgGlyAlaAlaGlyAlaGluValAlaVal 19
Db 648 ATGGGAAGTAAACATGCCAGCAAGTCCAGGGGACCGCGGGCGAGAGGTCCCGCGTTC 589
QY 20 GluValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaThrGlyGly 39
Db 588 GAGTCACGCAGATCGTCGGCGTCCGGACGAGATCCAGATCCCGCGGCGACCGGGGT 529
QY 40 ---ValAlaLysValAlaProArgArgLysArgAla-----ProAla 52
Db 528 GCTGTGGCGAAGATCGCACCGAGGAGGAAGAGCGCTGGCGCGCTCCACACGTGGCGCG 469
QY 53 GlyGluProAlaAlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeu 72
Db 468 GGGGAGCCT---GCTGCCGTCCGCGCTGGAGGGGACGGCGCGCTGCTACATCTGTG 412
QY 73 ArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAspSerValPro 92
Db 411 CGCAGCCGATGCTGTTCATGGCACCGCTTCAGCGCGAGCTG-----TCGCTTCCG 361
QY 93 ThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaGly 112
Db 360 ACGCCGCGGAGGTGCT-----GGTGCAGGACAGCAGGCGCGCGCTCGCGCGCGGG 307
QY 113 LeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyGlyGlnArgGly 132
Db 306 CTCTCGGTTGCTCCAGCACGGCGTCTGCTGGTGGACGTCCGG---GGTCATCAGCGCAGG 250
QY 133 SerHisThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuValAsp 152
```

```
|||||
249 AGCCACGCTGCCGTCC---GACGCTGCGGAGGTTGGCGGGATCACGCC---CCGGAT 196
|||||
153 ValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThrPro 172
|||||
195 GTCTCG---GCGAGCAACTCGGGGAGCGGCCAGACCGCGAGAGGAGAGACGCCA 139
|||||
173 SerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHisLys--- 191
|||||
138 TCGAGCCGGCGGCACGCGAGCTCAGCGATCTGGAGTCGGATCTGGCGGGCACAAGACT 79
|||||
192 ThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeuIleValProProAlaHis 211
|||||
78 ACTGCCCGGTCTGTCGGCGGCAACGTCGGCTGCGCAGCCGATCGTGTCCACCAACA 19
|||||
212 GluIleGlnGluPhePhe 217
|||||
18 GAGATCGAGGAGTTCTTC 1
```

```
RESULT 4
CA106521
LOCUS
DEFINITION
SCQHR1013C07.g HR1 Saccharum officinarum cDNA clone SCQHR1013C07
5', mRNA sequence.
CA106521
CA106521.1 GI:34959828
EST.
Saccharum officinarum
Saccharum officinarum
```

705 bp mRNA linear EST 23-SEP-2003

```
CA106521
SCQHR1013C07.g HR1 Saccharum officinarum cDNA clone SCQHR1013C07
5', mRNA sequence.
CA106521
CA106521.1 GI:34959828
EST.
Saccharum officinarum
Saccharum officinarum
```

CA106521.1 GI:34959828

EST.

Saccharum officinarum

Saccharum officinarum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum.

1 (bases 1 to 705)

Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.

The libraries that made SUCEST

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found

through the Brazilian Clone Collection Center (BCCC) at

<http://www.bcccenter.fcav.unesp.br>

Plate: 013 row: C column: 07

Seq primer: T7 Promoter Primer.

Location/Qualifiers

1.705

/organism="Saccharum officinarum"

/mol_type="mRNA"

/db_xref="taxon:4547"

/clone="SCQHR1013C07"

/lab_host="DH10B"

/clone_lib="HR1"

/note="Organ: seedlings inoculated with Herbaspirillum

rubrisubalbicans; Vector: pSport1; Site 1: SalI; Site 2:

NotI; An unidirectional cDNA library generated from

[seedlings inoculated with Herbaspirillum

rubrisubalbicans]. cDNA was prepared from polyA+ mRNA

using Superscript Plasmid System Kit (Invitrogen). The

double-strand cDNAs were fractionated in a sepharose

CL-2B 40cm-columns and fragments sizing between 0.8 and

1.5 Kb were directionally cloned into the vector. Details

of each source of RNA and library construction can be

obtained at <http://sucest.lad.ic.unicamp.br/public>"

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

1.67e-35

690.50

81.95%

Length:

Matches:

Conservative: 6

Best Local Similarity: 79.02% Mismatches: 18
Query Match: 52.95% Indels: 19
DB: 13 Gaps: 11

US-09-993-808B-2 (1-256) x CA106521 (1-705)

QY 1 MetGlyLysTyrMetArg---LysCysArgGlyAlaAlaGlyAlaGluValAlaVal 19

Db 113 ATGGGAAGTACATCGCAGCAAGTGCAGGGGCGCGGAGGAGGAGTCCGCGCGT 172

QY 20 GluValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaThr---Gly 38

Db 173 GAGTCAACGAGTCTCGGCGTCCGACGAGGTCCAGGTCCGCGGCGCGGCGG 232

QY 39 GlyValAlaLysValAlaProArgArgLysArgAla-----ProAla 52

Db 233 GGTGTGGCAAGTCCGACCGAGGAGGAAGAGGCGTGGCGCGCCCAACGTGGCG 292

QY 53 GlyGluProAlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeu 72

Db 293 GGGGAGCCT---GCTGCCGTGCGCGCTGGAGGGGACGCGGGAAGTGTCTACATCTG 349

QY 73 ArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAspSerValPro 92

Db 350 CGCAGCCGATGCTGTTCATGGCACCGCCTCAGCCGACGCG---TCGCTCCG 400

QY 93 ThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaGly 112

Db 401 ACGCCGGCGGAGGCTGCT-----GGTGCAGGACANACAGGCGCGGCGTCCGCGCG 454

QY 113 LeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlnArgGly 132

Db 455 CTCTCGGTNGTCTCCAGCACGCGTCTCGGTGGACGTCGGG---GGTCATCAGCGCAAG 511

QY 133 SerHisThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuValAsp 152

Db 512 AGCCACGCTGCCGCTCC---GACGCTGCGGAGGTGGCGGGGATCACGCC---CCGGAT 565

QY 153 ValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThrPro 172

Db 566 GTCTCN---GCGAGCAACTCGGGGAGCGGCCCAACCGCAGAGAGAGAGACGCC 622

QY 173 SerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHisLys--- 191

Db 623 TCGAGCGGGCGCACGCGAGTCAACGATCTGGAGTTCGATCCGCGGGGACAAAGATT 682

QY 192 ThrGlyProSerLeu 196

Db 683 ACGGCCCGGTCTGTTG 697

FEATURES

source

RESULT 5

CA256119

LOCUS

DEFINITION

5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Saccharum officinarum

Saccharum officinarum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Saccharum.

1 (bases 1 to 661)

Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.

The libraries that made SUCEST

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

CA256119 661 bp mRNA linear EST 26-SEP-2003
SCJFL4187B06.g FL4 Saccharum officinarum cDNA clone SCJFL4187B06

Email: parruda@unicamp.br
 Clone distribution: clone distribution information can be found
 through the Brazilian Clone Collection Center (BCCC) at
 http://www.bccccenter.fcav.unesp.br
 Plate: 187 row: B column: 06
 Seq primer: T7 Promoter Primer.
 Location/Qualifiers

FEATURES

1..661
 /organism="Saccharum officinarum"
 /mol_type="mRNA"
 /db_xref="taxon:4547"
 /clone="SCJFL4187B06"
 /lab_host="DH10B"
 /clone_lib="FL4"

/note="Organ: Developed inflorescence and rachis
 (20cm-long); Vector: pSport1; Site 1: Sali; Site 2: NotI;
 An unidirectional cDNA library generated from [Developed
 inflorescence and rachis (20cm-long)]. cDNA was prepared
 from polyA+ mRNA using SuperScript Plasmid System Kit
 (Invitrogen). The double-strand cDNAs were fractionated
 in a sepharose CL-2B 40cm-columns and fragments sizing
 between 0.8 and 1.5 Kb were directionally cloned into the
 vector. Details of each source of RNA and library
 construction can be obtained at
 http://sucest.fad.ic.unicamp.br/public"

ORIGIN

Alignment Scores:
 Pred. No.: 1.5e-33 Length: 661
 Score: 660.00 Matches: 151
 Percent Similarity: 83.16% Conservativity: 7
 Best Local Similarity: 79.47% Mismatches: 14
 Query Match: 50.61% Indels: 18
 DB: 14 Gaps: 10

US-09-993-808B-2 (1-256) x CA256119 (1-661)

QY 1 MetGlyLysTyrMetArg---LysCysArgGlyAlaAlaGlyAlaGluValAlaAlaVal 19
 Db 121 ATGGGAAGTACATCGCAGCAAGTGCAGGGGCGCCCGGGGCGAGGAGTCCGCCCGTC 180
 QY 20 GluValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThr---Gly 38
 Db 181 GAGGTACGCAGGTGTCGGCGTCCGGACGAGGTCCAGTCCGGCGGCGGCGGCGGT 240
 QY 39 GlyValAlaLysValAlaProArgArgLysArgAla-----ProAla 52
 Db 241 GGTGTGGCGAAGTTCGACCCGAGGAGGAAGGCGTGGCGCCCGCCCGCAACGTGGCGGCG 300
 QY 53 GlyGluProAlaAlaValSerAlaGlyGlyGlyGlyGlySerCysTyrIleHisLeu 72
 Db 301 GGGGAGCCT---GCTGCCGTCGGCGTGGAGGAGCGGCGGGAAGTCTACATCCATCTG 357
 QY 73 ArgSerArgMetLeuPheMetAlaProGlnProGlnProSerValAspSerValPro 92
 Db 358 CGCAGCCGCATGCTGTTTCATGGCACCGCCTCAGCCGCGACCG-----TCGCTCCG 408
 QY 93 ThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaGly 112
 Db 409 ACGCCGGCGGAGGCTGCT-----GGTGACGAGCAGCAGGGCGCGGCTCGCGGCGGG 462
 QY 113 LeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlyGlnArgGly 132
 Db 463 CTCTCGGTTGCTCCAGCAGCGCTCGTCCGTGGACGTCCGG---GGTCATCAGCGCAGG 519
 QY 133 SerHisThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuValAsp 152
 Db 520 AGCCACGCCTGCCGCTCC---GACGCTCGGAGGTTGGCGGGGATCAGGCC---CCGGAT 573
 QY 153 ValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThrPro 172
 Db 574 GTCTCG---GCGAGCAACTCGGGGAGCGGGCCAGACCGCGAAGGAGAAACGACGCCA 630

QY 173 SerSerArgAlaHisGlyGluLeuSerAsp 182
 Db 631 TTTTAACCGCGCACGCGCGAGCTTAACGAT 660
 RESULT 6
 CC610974/c
 LOCUS CC610974 666 bp DNA linear GSS 18-JUN-2003
 DEFINITION OGVFR67TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0551L14,
 genomic survey sequence.
 ACCESSION CC610974
 VERSION CC610974.1 GI:31972395
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
 clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 666)
 AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
 Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
 TITLE Consortium for Maize Genomics
 JOURNAL Unpublished (2002)
 COMMENT Other GSSs: OGVFR67TH
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TF
 Class: sheared ends.
 FEATURES
 Location/Qualifiers
 1..666
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBMA0551L14"
 /clone_lib="ZM_0.7_1.5_KB"
 /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN

Alignment Scores:
 Pred. No.: 2.76e-33 Length: 666
 Score: 656.00 Matches: 142
 Percent Similarity: 81.14% Conservativity: 0
 Best Local Similarity: 81.14% Mismatches: 7
 Query Match: 50.31% Indels: 26
 DB: 29 Gaps: 5
 US-09-993-808B-2 (1-256) x CC610974 (1-666)
 QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
 Db 467 ATGGGAAGTACATCGCAAGTGCAGGGGCGCGGCGGAGGTCGCCCGCGTCGAG 408
 QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal 40
 Db 407 GTTACGACGGTCTCGCGTCCGACGAGGTCCAGTCCGGCGGCGGAGCCCTGTCGCGCTGTC 348
 QY 41 AlaLys---ValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaVal 59
 Db 347 GCGAAGTTCGTCGCCCGCAGGAGGAAGAGGGCGCGCGGGGAGCCTGTCGCGCGCTGTC 288
 QY 60 SerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMet 79
 Db 287 GCGCTGGTGGGACGCGGGAAGTCTACATCCACCTGCGAAGCCGATGCTGTTTCATG 228
 QY 80 AlaProGlnProGlnProSerValAspSerValProThrPro---ValGluAlaAla 98
 Db 227 GCACCGCCTCAGCCCGCAGCCG-----TCGGTCCGACCCCGGCGGAGGCTGCT 177


```
QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyValAlaGluValAlaAlaValGlu 20
Db 146 ATGGGAAGTACATGCGCAAGTGCAGGGCGCGCAGCGCGGAGGTGCGCCCGCTCGAG 205
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal 40
Db 206 GTTACGCAGGTGCTCGGCGTCCGACGAGGTCCAGGTCCGCGCGCGGACCGCGGCTGC 265
QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaAlaValSer 60
Db 266 GCGAAGTCCGCCCGCAGGAGAGAGGGCGCGCGGGGGAGCCTGCTGCCCGCTGAGC 325
QY 61 AlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAla 80
Db 326 GCTGTGGGGACGGCGGAGTGTACTACATCCACCTGCTAGCCGCATGCTGTTTCATGGCA 385
QY 81 ProProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAspGly 100
Db 386 CCGCCTCAGCCGACGCCGTCGTTGACTCGGTTCCGACCCCGGTGGAGGCTGCTGATGCC 445
QY 101 AlaAlaGlyGlnGlnGlyAlaAlaLeuAlaGlyLeuSerArgCysSerSerThrAla 120
Db 446 GCTGCAGGACAGCAGGGCGCGGCTCGCTGCCGGGCTCTCGGTTGCTCCAGCAGGCG 505
QY 121 SerSerValAsnLeuGlyLeu 127
Db 506 TCGTCTGTGAACCTTGGGCTTG 526
```

```
RESULT 9
CC440801/c
LOCUS PUHLN53TB ZM_0.6_1.0_KB Zea mays genomic clone ZMBTa478J10,
DEFINITION genomic survey sequence.
```

```
ACCESSION CC440801
VERSION CC440801.1 GI:30942238
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
```

```
REFERENCE 1 (bases 1 to 838)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other GSSs: PUHLN53TD
Contact: Cathy Whitelaw
TIGR
```

```
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
```

```
FEATURES
source Location/Qualifiers
1..838
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBTa478J10"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
CoT selected genomic DNA library"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 6.93e-29 Length: 838
Score: 590.50 Matches: 123
Percent Similarity: 94.03% Conservatives: 3
Best Local Similarity: 91.79% Mismatches: 3
```

```
Query Match: 45.28% Indels: 5
DB: 28 Gaps: 3
US-09-993-808B-2 (1-256) x CC440801 (1-838)
QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyValAlaGluValAlaAlaValGlu 20
Db 390 ATGGGAAGTACATGCGCAAGTGCAGGGCGCGCAGCGCGGAGGTGCGCCCGCTCGAG 331
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal 40
Db 330 GTTACGCAGGTGCTCGGCGTCCGACGAGGTCCAGGTCCGCGCGCGGACCGCGGCTGC 271
QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaAlaVal 59
Db 270 GCGAAGTCCGCCCGCAGGAGAGAGGGCGCGCGGGGGAGCCTGCTGCCCTCGGTG 211
QY 60 SerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMet 79
Db 210 GCGCTGTGGGACGGCGGAGTGTACTACATCCACCTGCTAGCCGCATGCTGTTTCATG 151
QY 80 AlaProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAsp 99
Db 150 GCACCGCCTCAGCCGCGAGCGG-----CCGTCGGTTCCGACCCCGCGGAGGCTGCTGAT 97
QY 100 GlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaGlyLeuSerArgCysSerSerThr 119
Db 96 GCGCTGCAGGACAGCAGGGCGCGGCTCGCGGCGCGGCTCTCGGTTGCTCCAGCAGC 37
QY 120 AlaSerSerValAsnLeuGlyLeuGlyGlnArgGlySer 133
Db 36 GCGTCGTCGGTGAC-----GTGGGGGGTTCAGCGCGGGAGC 1
```

```
RESULT 10
CG136784
LOCUS PUIDC25TB ZM_0.6_1.0_KB Zea mays genomic clone ZMBTa0556E01,
DEFINITION genomic survey sequence.
```

```
ACCESSION CG136784
VERSION CG136784.1 GI:34026526
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
```

```
REFERENCE 1 (bases 1 to 645)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other GSSs: PUIDC25TD
Contact: Cathy Whitelaw
TIGR
```

```
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
```

```
FEATURES
source Location/Qualifiers
1..645
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBTa0556E01"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
CoT selected genomic DNA library"
```

ORIGIN


```
Db      505 TCCGCGAGCGACTCGCGAGCGCGCGTGACCGGAGAGAGAGAAACAACCTCCATCAAGC 564
QY      175 ArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAla---GlyHisLysThrGly 193
Db      565 TTTCTCCCGGAGAGGTGAGCGATCTGGAGTCGATCTGGCTGGAGACAGAAGCGCAGC 624
QY      194 ProSerLeuPro---AlaAlaThrProAlaAlaGluLeuIle-----ValPro 208
Db      625 CGTCCACTACCTTCTCGGCAACAGCCTCAGCAGACAGCAAGCCACGCGGCCGAAGATTCCG 684
QY      209 ProAlaHisGluIleGlnGluPhePheAlaAlaGluAlaAlaGlnAlaLysArgPhe 228
Db      685 CCGGCGCGGAGATCAGGGCTTCTTCGGCGCGCGCGCAAGAGGCTGAGGCCAAGCGCTTC 744
QY      229 AlaSerLys 231
Db      745 GCCGCCAAG 753

RESULT 12
AY110319
LOCUS   AY110319                      707 bp      mRNA      linear      HTC 17-OCT-2002
DEFINITION   Zea mays CL528_-1 mRNA sequence.
ACCESSION   AY110319
VERSION     AY110319.1  GI:21214635
KEYWORDS     HTC.
SOURCE      Zea mays
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 707)
AUTHORS     Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
            Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE       Maize Mapping Project/DuPont Consensus Sequences for Design of
            Overgo Probes
JOURNAL     Unpublished (2002)
REFERENCE   2 (bases 1 to 707)
AUTHORS     Coe,E.H.
TITLE       Direct Submision
JOURNAL     Submitted (25-APR-2002) Maize Mapping Project, University of
            Missouri, Columbia, MO 65211, USA
COMMENT     If you are interested in getting corresponding physical clones,
            these are publicly available from ZmDB and may be found by BLAST
            searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
            www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
            maize cDNA sequences is either Virginia Walbot, Stanford or Pat
            Schnable, Iowa State, then clones may be requested from ZmDB:
            www.zmdb.iastate.edu.
FEATURES    Location/Qualifiers
            1..707
               /organism="Zea mays"
               /mol_type="mRNA"
               /db_xref="MaizeDB:632312"
               /db_xref="taxon:4577"
               /clone_lib="Maize Mapping Project/DuPont Consensus
               Library"
               /note="this sequence is part of a project of EST
               assemblies resulting from the application of public
               contigs to seed DuPont contigs; this resource was
               assembled by DuPont as part of a collaboration for the
               overgo addressing of BACs in conjunction with the Maize
               Mapping Project"

ORIGIN
Alignment Scores:
Pred. No.:      3,46e-27      Length:      707
Score:          563.00      Matches:    110
Percent Similarity: 88.80%      Conservative: 1
Best Local Similarity: 88.00%      Mismatches:  14
Query Match:    43.17%      Indels:     0
DB:             11      Gaps:         0

US-09-993-808B-2 (1-256) x AY110319 (1-707)
```

```
QY      132 GlySerHisThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuVal 151
Db      10  GGGAGCCACACCTGCGCTCCGACGACGCTGCAGAGGCTGGCGGGATCACGTCCTGGTG 69
QY      152 AspValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThr 171
Db      70  GTGGATGTCTCGGCGAGCAACTCCGGAGCGGCCACAGACCGGAGAGAGACGACG 129
QY      172 ProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHisLys 191
Db      130 CCATCGAGCCGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGATCTGGCGGGCACAAG 189
QY      192 ThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeuIleValProAlaHis 211
Db      190 ACTGGCCCGTCTGCTACCGCGGCAACGCCGCTCGGAGCTGATCGTCCCGCCAGCACAC 249
QY      212 GluIleGlnGluPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAlaSerLys 231
Db      250 GAGATCCAGGAGTTCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 309
QY      232 TyrAsnPheAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGluTrpAla 251
Db      310 TACAACCTTCGACTTCGTCCGCGCGCTGCCCTCCAGCGCGCGCGGTCGAGTGGCGG 369
QY      252 ProValValSerIle 256
Db      370 CCGGTGTCAGCATC 384

RESULT 13
CA104568
LOCUS     CA104568                      585 bp      mRNA      linear      EST 23-SEP-2003
DEFINITION   SCJFHR1034B05.g HR1 Saccharum officinarum cDNA clone SCJFHR1034B05
            5', mRNA sequence.
ACCESSION   CA104568
VERSION     CA104568.1  GI:34957875
KEYWORDS     EST.
SOURCE      Saccharum officinarum
ORGANISM    Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum.
REFERENCE   1 (bases 1 to 585)
AUTHORS     Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLE       The libraries that made SUCEST
JOURNAL     Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 034 row: B column: 05
            Seq primer: T7 Promoter Primer.
FEATURES    Location/Qualifiers
            1..585
               /organism="Saccharum officinarum"
               /mol_type="mRNA"
               /db_xref="taxon:4547"
               /clone="SCJFHR1034B05"
               /lab_host="DH10B"
               /clone_lib="HR1"
               /note="Organ: seedlings inoculated with Herbaspirillum
               rubrisubalbicans; Vector: pSport1; Site_1: SalI; Site_2:
               NotI; An unidirectional cDNA library generated from
               [seedlings inoculated with Herbaspirillum
               rubrisubalbicans]. cDNA was prepared from polyA+ mRNA
               using Superscript Plasmid System Kit (Invitrogen). The
               double-strand cDNAs were fractionated in a sepharose
```


CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public>

ORIGIN

Alignment Scores:
Pred. No.: 1.28e-24 Length: 585
Score: 522.00 Matches: 128
Percent Similarity: 79.04% Conservative: 4
Best Local Similarity: 76.65% Mismatches: 17
Query Match: 40.03% Indels: 18
DB: 13 Gaps: 10

US-09-993-808B-2 (1-256) x CA104568 (1-585)

QY 1 MetGlyLysTyrMetArg---LysCysArgGlyAlaAlaGlyAlaGluValAlaAlaVal 19
Db 113 ATGGGAAGTACATGCGCAGCAAGTGCAGGGGCGCGCGGAGGAGTCCGCCCGCGTC 172
QY 20 GluValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThr---Gly 38
Db 173 GAGGTACAGCAGTCTCGCGTCCGACGAGTCCAGGTCCGCGCGCGGCGACCGGGGT 232
QY 39 GlyValAlaLysValAlaProArgArgLysArgAla-----ProAla 52
Db 233 GGTGTGGCGAAGTCTGCACCGAGGAGGAGGAGGCGCTGGCGCGCGCAACGTGGCGCG 292
QY 53 GlyGluProAlaAlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeu 72
Db 293 GGGGAGCCT---GCTGCCGTGGCGCTGGAGGGACGGCGGAGCTGCTACATCCATCTG 349
QY 73 ArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAspSerValPro 92
Db 350 CGCAGCCGCTGCTGCTCATGGCACCGGCTCAGCGCGAGCCG-----TCGCTTCG 400
QY 93 ThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGly 112
Db 401 ACGCCGGCGGAGGTGCT-----GGTGCANGACAGCAGGGCGCGGCTCGCGGACGG 454
QY 113 LeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlnArgGly 132
Db 455 CTCTCGGTTGCTCCAGCTAGGAGTCNTCGGTGACGTGCGG---GGTCATCAACGCAG 511
QY 133 SerHisThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuValAsp 152
Db 512 AGCCACGCCTGACGCTCC---GACGCTGCGGAGGTGGCGGGGATCACGCC---CCGGAT 565
QY 153 ValSerAlaAlaSerAsnSer 159
Db 566 GTCTCG---GCGAGCAACTCG 583

RESULT 14
CA104689/c
LOCUS CA104689 642 bp mRNA linear EST 23-SEP-2003
DEFINITION SCUJFHR1034G09.b HR1 Saccharum officinarum cDNA clone SCUJFHR1034G09 3', mRNA sequence.

ACCESSION CA104689.1 GI:34957996
VERSION CA104689
KEYWORDS EST.

SOURCE Saccharum officinarum
ORGANISM Saccharum officinarum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum.

REFERENCE 1 (bases 1 to 642)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
CONTACT: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found through the Brazilian Clone Collection Center (BCCC) at <http://www.bcccenter.fcav.unesp.br>
Plate: 034 row: G column: 09
Seq primer: SP6 Promoter primer.
Location/Qualifiers

FEATURES
source

1. .642
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCJFHR1034G09"
/lab_host="DH10B"
/clone_lib="HR1"

/note="Organ: seedlings inoculated with Herbaspirillum rubrisubalbicans; Vector: pSport1; Site_1: Sali; Site_2: Not1; An unidirectional cDNA library generated from [seedlings inoculated with Herbaspirillum rubrisubalbicans]. cDNA was prepared from polyA+ mRNA using Superscript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public>"

ORIGIN

Alignment Scores:
Pred. No.: 2.63e-24 Length: 642
Score: 518.00 Matches: 132
Percent Similarity: 69.12% Conservative: 18
Best Local Similarity: 60.83% Mismatches: 49
Query Match: 39.72% Indels: 19
DB: 13 Gaps: 10

US-09-993-808B-2 (1-256) x CA104689 (1-642)

QY 9 ArgGlyAlaAlaGlyAlaGluValAlaAlaValGluValThrGlnValValGlyValArg 28
Db 620 AGGCCCGCTTCGGTTGAGGAGATCACCCGCAATCGAGTCAACCAGAT-GTCGGCGTCCGT 562
QY 29 ThrArgSerArgSerAlaAlaAlaThr---GlyGlyValAlaLysValAlaProArgArg 47
Db 561 ACAGATATAGTTCAGACGCTGCAACTGGGGGTGGTCTGCCGAAGTTCGACCGAGGAGG 502
QY 48 LysArgAla-----ProAlaGlyGluProAlaAlaAlaValSerAla 61
Db 501 AAGAAGCGCTGACACCGCAACAGATGACACGGGGGAACCT---ACTGCGATCGGCAAT 445
QY 62 GlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAlaPro 81
Db 444 GGAGGGTACGTCGGAAGCTGCTACATCCATATCGCAGCGGCGCATGGTGTTCATGTACCG 385
QY 82 ProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAspGlyAla 101
Db 384 CATCAGCAGCAG-----CAGTGGCTTCAGACGCGCGGAGCCTGCT-----GGT 340
QY 102 AlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThrAlaSer 121
Db 339 GCAGAACAGCAGCGCGCGGCTAGCGCCCGGACTATCGGTTGCTCCAGCAGGAGTCG 280
QY 122 SerValAsnLeuGlyLeuGlyGlyGlnArgGlySerHisThrCysArgSerTyrAspAla 141
Db 279 TCGGTGGAGATCGGG---GGTCATCAGCGCAGGAGTACGCTAGCCATCA---GACACT 226
QY 142 AlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGlySer 161
Db 225 CCGAGGTTGACGGGATCACGCC---CCGGAAGTTTCG---GCGAGCAAGTCGGGGAGC 172
QY 162 GlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyGluLeuSer 181

```
Db 171 GGCCAGACCGGAGGAGAGAGACGACGCCATCGAGCCGGGCGCAGCGGAGATCAAC 112
QY 182 AspLeuGluSerAspLeuAlaGlyHisLys---ThrGlyProSerLeuProAlaAlaThr 200
Db 111 GATATGGAGTCGATTGGCGGGGCACAAAGATACTGCCCGGTAGTAGCCGGCGCAACG 52
QY 201 ProAlaAlaGluLeuValProProAlaHisGluIleGlnGluPhePhe 217
Db 51 TCGGTCGCGAGCCGATCGTGCCACCAGCACAAAGAGATCGAGGAGTTCTTC 1

RESULT 15
CC729680 644 bp DNA linear GSS 23-JUN-2003
OGUBY35TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma0404F21,
genomic survey sequence.
ACCESSION CC729680
VERSION CC729680.1 GI:32148613
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE 1 (bases 1 to 644)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGUBY35TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
source
Location/Qualifiers
1..644
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0404F21"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Alignment Scores:
Pred. No.: 2.85e-24 Length: 644
Score: 517.50 Matches: 128
Percent Similarity: 59.17% Conservative: 1
Best Local Similarity: 58.72% Mismatches: 7
Query Match: 39.69% Indels: 83
DB: 29 Gaps: 4

US-09-993-808B-2 (1-256) x CC729680 (1-644)
QY 82 ProGlnProGlnProSerValAspSerValProThrPro---ValGluAlaAlaAspGly 100
Db 3 CCTCAGCGCGAGCCG-----TCGGTTCCGACCCCGCGCGGAGGCTGCTGATGGC 53
QY 101 AlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThrAla 120
Db 54 GCTGCAGGACAGCAGCGCGGCTGCTCGCGGCCCGGGCTCTCGCGTGTCTCCAGCAGCGG 113
QY 121 SerSerValAsnLeuGlyLeuGlyGlyGlnArgGlySerHisThrCysArgSerTyrAsp 140
Db 114 TCGTCGGTGAACCTTGGGTTGGGGGTTCAGCGCGGAGCCACACCTGCCGCTCCGACGAC 173
QY 141 AlaAlaGluAla-----144
```

```
Db 174 GCTGCAGAGGT-GAGAGAGAGAGAGAGAGCCAAACCCCTAGAATTATTTCGAATTCAAA 232
QY 144 -----144
Db 233 ACGGACCGAGATCTCCAGCGCTACTCGTTGCTGAATTTCCGTGCGCTTTCTTTTGGTTTGG 292
QY 145 -----GlyGlyAspHisValLeuValAspValSerAlaAlaSer 157
Db 293 GTCTCTACGACGTAAACAGGCTGGCGGGGATCACGTCCTGTTGGTGGATGTCTCGGCGAGC 352
QY 158 AsnSerGlySerGlyProAspArgGlu-----166
Db 353 AACTCCGGAGCGGCCCCAGACCGGAGAGGTTGGAGAATTCTCCCTTTGTGTCTCGCGAT 412
QY 166 -----166
Db 413 TCCTGCTGAATTATGCCTTCGCGCGTGTCTCCTGACGACGTTCCGCGCCTGTGGATTTTGG 472
QY 167 -----ArgArgGluThrThrProSerSerArgAlaHisGlyGluLeuSerAspLeuG1 184
Db 473 AAAATTTCAGGAGAGACGACGCCATCGAGCCGGCGGCACGGCGAGCTCAGCGATCTGGA 532
QY 184 uSerAspLeuAla-GlyHisLysThrGlyProSerLeuProAlaAlaThrProAlaAlaG 204
Db 533 GTCGGATCTGGCGGGGGCACAAAGACTGGCCCGTCCGCTACCGCGGCAACGCCGCGTCCGG 592
QY 204 luLeuIleValProProAlaHisGluIleGlnGluPhePheAlaAla 220
Db 593 AGCTGATCGTGC CGCCAGACACAGATCCAGGAGTTCTTCGCCCGCGGCC 642
```

Search completed: October 2, 2004, 11:51:10
Job time : 2478 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2004, 08:37:55 ; Search time 3090 Seconds
(without alignments)
3590.879 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 1304
Sequence: 1 MGKYMRCRGAAGAEVAAVE.....VRGVPLDAGGRFEWAPVVS I 256

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delopt 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09933808/runat_01102004_171702_13026/app_query.fasta_1.455
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09933808@cgn 1 1 5600 @runat_01102004_171702_13026 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB | ID | Description |
|------------|-------|---------------|--------|----|-----------|--------------------|
| 1 | 669 | 51.3 | 1242 | 6 | AX406686 | AX406686 Sequence |
| 2 | 669 | 51.3 | 1266 | 8 | AK103084 | AK103084 Oryza sat |
| C 3 | 569 | 43.6 | 135670 | 8 | AP005002 | AP005002 Oryza sat |
| 4 | 399.5 | 30.6 | 568 | 6 | AX406682 | AX406682 Sequence |
| 5 | 279 | 21.4 | 681 | 6 | AX406731 | AX406731 Sequence |
| 6 | 242 | 18.6 | 1073 | 6 | AX406733 | AX406733 Sequence |
| C 7 | 242 | 18.6 | 151208 | 2 | AP003525 | AP003525 Oryza sat |
| C 8 | 242 | 18.6 | 187916 | 2 | AP003458 | AP003458 Oryza sat |
| 9 | 240.5 | 18.4 | 1045 | 8 | AK064723 | AK064723 Oryza sat |
| 10 | 238 | 18.3 | 1027 | 8 | AK073804 | AK073804 Oryza sat |
| 11 | 235 | 18.0 | 1067 | 6 | AX406720 | AX406720 Sequence |
| 12 | 215.5 | 16.5 | 797 | 6 | AX463079 | AX463079 Sequence |
| 13 | 212.5 | 16.3 | 669 | 8 | ATH301554 | AJ301554 Arabidops |
| 14 | 207 | 15.9 | 962 | 8 | LES441250 | AJ441250 Lycopersi |
| 15 | 199.5 | 15.3 | 804 | 6 | AX008805 | AX008805 Sequence |
| 16 | 199.5 | 15.3 | 804 | 6 | AX463073 | AX463073 Sequence |
| 17 | 199.5 | 15.3 | 804 | 8 | CRJ002173 | AJ002173 Chenopodi |
| 18 | 190.5 | 14.6 | 875 | 6 | A98056 | A98056 Sequence 3 |
| 19 | 190.5 | 14.6 | 875 | 6 | BD075688 | BD075688 Cyclin-de |
| 20 | 179.5 | 13.8 | 861 | 8 | ATH301555 | AJ301555 Arabidops |
| 21 | 178 | 13.7 | 870 | 8 | AF123315 | AF123315 Arabidops |
| 22 | 178 | 13.7 | 1290 | 6 | AX463075 | AX463075 Sequence |
| 23 | 177.5 | 13.6 | 1312 | 8 | AK117586 | AK117586 Arabidops |
| 24 | 170.5 | 13.1 | 927 | 8 | AB029483 | AB029483 Pisum sat |
| 25 | 169.5 | 13.0 | 1014 | 8 | LES441249 | AJ441249 Lycopersi |
| 26 | 163 | 12.5 | 1072 | 8 | NTA517189 | AJ517189 Nicotiana |
| 27 | 159.5 | 12.2 | 789 | 8 | NTO297906 | AJ297906 Nicotiana |
| 28 | 153 | 11.7 | 570 | 8 | ATH301556 | AJ301556 Arabidops |
| C 29 | 152.5 | 11.7 | 15311 | 1 | AE007164 | AE007164 Mycobacte |
| 30 | 152 | 11.7 | 2051 | 8 | AK107643 | AK107643 Oryza sat |
| 31 | 152 | 11.7 | 105659 | 2 | AP003961 | AP003961 Oryza sat |
| 32 | 152 | 11.7 | 162926 | 2 | AP005386 | AP005386 Oryza sat |
| 33 | 152 | 11.7 | 207823 | 2 | AC145542 | AC145542 Oryctolag |
| 34 | 151.5 | 11.6 | 493 | 6 | AX406678 | AX406678 Sequence |
| C 35 | 151.5 | 11.6 | 178743 | 10 | AL808102 | AL808102 Mouse DNA |
| 36 | 150.5 | 11.5 | 168560 | 8 | AC118132 | AC118132 Oryza sat |
| 37 | 149.5 | 11.5 | 154494 | 2 | AC120885 | AC120885 Oryza sat |
| C 38 | 149.5 | 11.5 | 316050 | 1 | BX248346 | BX248346 Mycobacte |
| 39 | 149 | 11.4 | 984 | 6 | AX412881 | AX412881 Sequence |
| 40 | 149 | 11.4 | 984 | 6 | AX505332 | AX505332 Sequence |
| 41 | 148.5 | 11.4 | 109081 | 9 | AL358074 | AL358074 Human DNA |
| C 42 | 145.5 | 11.2 | 7467 | 8 | AY186602 | AY186602 Chlamydom |
| 43 | 145 | 11.1 | 533 | 6 | AX008799 | AX008799 Sequence |
| 44 | 145 | 11.1 | 533 | 6 | AX463068 | AX463068 Sequence |
| C 45 | 143.5 | 11.0 | 2285 | 9 | BC040188 | BC040188 Homo sapi |

ALIGNMENTS

AX406686
LOCUS AX406686 1242 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 9 from Patent WO0228893.
ACCESSION AX406686
VERSION AX406686.1 GI:21439634
KEYWORDS Oryza sativa
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartioideae; Oryzeae; Oryza.
1
REFERENCE Droual, A.-M., Frankard, V.M.-N., Inze, D., Mironov, V., Peres, B.A.D.
AUTHORS and Hatzfeld, Y.
TITLE Plant cyclin-dependent kinase inhibitors
JOURNAL Patent: WO 0228893-A 9 11-APR-2002;
CROPDESIGN N V (BE)
FEATURES
source Location/Qualifiers
1..1242
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"
ORIGIN
Alignment Scores:
Pred. No.: 4.6e-22 Length: 1242
Score: 669.00 Matches: 161
Percent Similarity: 67.42% Conservative: 19
Best Local Similarity: 60.30% Mismatches: 67
Query Match: 51.30% Indels: 20
DB: 6 Gaps: 10
US-09-993-808b-2 (1-256) x AX406686 (1-1242)
QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaGlyAlaGluValAlaAlaValGlu 20
Db 18 ATGGCAAGTACATGAGGAAGTTCAGGGGGCCACGGGGAGGAGTTGGCCGCATGGAG 77
QY 21 ValThrGlnValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGly--- 39
Db 78 GTCACGAGGTGTTGGCTCCGACGAGTTCGAGTCCGAGCGCGCGCGCGCGCGCAGG 137
QY 40 -----ValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGlu 54
Db 138 ACACGAAGTGCAGGGCGGCTCGGGCGCTCCACCAGGAGGAGGAGGCGCTGCTGCCG 197
QY 55 ProAlaAlaAlaValSerAlaGlyGlyAspGlySerCysTyrIleHisLeuArgSer 74
Db 198 ACGGCGGTCTGTGGGACTACTCGCGTACGGGGGAGCTGCTACCTCCAGCTGAGGAGC 257
QY 75 ArgMetLeuPheMetAlaProGlnProGlnProSerValAspSerValProThrPro 94
Db 258 CGCATGCTGTTTCATGGCCCCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 314
QY 95 ValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSer 114
Db 315 GCGGAGGCGCG---GGTTCGGGACGAGCGCGCGCGCATGCGCGGCTGGCCTCTCG 371
QY 115 ArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlyGlnArgGlySerHis 134
Db 372 CGTTGCTCCAGCACGGCGTCTCCGTGGAC-----GCGGCGGCTCAGGACGAGCCTC 425
QY 135 ThrCysArgSerThrAlaAlaGluAlaGlyGlyAspHisValLeuValAspValSer 154
Db 426 GCGTGGCGCTCC---GACGTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 476
QY 155 AlaAlaSerAsnSerGlySerGlyProAspArgGluArgArgGluThrThrProSerSer 174
Db 477 TCCGCGAGCGACTCGGCGAGCGCGCGCTGACCGCGAGGAGGAGGAGGAGGAGGAGG 536
QY 175 ArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAla---GlyHisLysThrGly 193
Db 537 TTTTCTCCCGCGAGGTGAGCGATCTGGAGTCTGGCTGGAGGAGGAGGAGGAGGAGG 596

QY 194 ProSerLeuPro---AlaAlaThrProAlaAlaGluLeuIle-----ValPro 208
Db 597 CGTCCACTACTTCTGCGGCAACAGCCTCAGCACGACGACGACGACGACGACGACG 656
QY 209 ProAlaHisGluIleGlnGluPhePheAlaAlaGluAlaAlaGlnAlaLysArgPhe 228
Db 657 CCGGCGCGGAGATCGAGCGGTTCTTCGCGCGCGCGGAGGAGGAGGAGGAGGAGG 716
QY 229 AlaSerLysTyrAsnPheAspPheValArgGlyValProLeuAspAlaGlyGlyArgPhe 248
Db 717 GCGGCCAAGTACAACTTCGACGCTCGTTCGCGCGCGTCCCTCGACGCC---GGTCGGTTC 773
QY 249 GluTrpAlaProValValSer 255
Db 774 GAGTGGACTCCGCTGGTTCAGC 794
AK103084 1266 bp mRNA linear PLN 24-JUL-2003
Oryza sativa (japonica cultivar-group) cDNA clone: J033118G18, full
insert sequence.
AK103084
AK103084.1 GI:32988293
FLI_CDNA; CAP trapper.
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartioideae; Oryzeae; Oryza.
1
REFERENCE The Rice Full-Length cDNA Consortium, National Institute of
AUTHORS Agrobiological Sciences Rice Full-Length cDNA Project Team,
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group: Ohtsuki, Y., Murakami, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M. and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)
22752273
12869764
2 (bases 1 to 1266)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kanagawa, S., Kato, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kondo, S., Konno, H., Kondo, H., Kouda, M.,
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,

ORIGIN /clone="P0486G03"

Alignment Scores:
Pred. No.: 9.08e-16 Length: 135670
Score: 569.00 Matches: 159
Percent Similarity: 46.04% Conservative: 21
Best Local Similarity: 40.66% Mismatches: 68
Query Match: 43.63% Indels: 143
DB: 8 Gaps: 12

US-09-993-808B-2 (1-256) x AP005002 (1-135670)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
Db 80195 ATGGCAAGTACATGAGGAAGTTCAGGGGGCCACGGGGGAGGAGTTGGCCGCCATGGAG 80136

QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGly--- 39
Db 80135 GTCACGAGGTGGTTGGCGTCGGACGAGGTCGAGGTCGGCAGCGGGGGCGGCGACG 80076

QY 40 -----ValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGlu 54
Db 80075 ACGACGAAGGTGAAGCGCGCTCGCGGCGTCCACGAGGAGGAAGCGCTGCTCCG 80016

QY 55 ProAlaAlaAlaValSerAlaGlyGlyAspGlySerCysTyrIleHisLeuArgSer 74
Db 80015 ACGGCGGTCTGTGGGACTACTCGCGGTGACGGCGGGAGTCTACCTCCAGCTGAGGAGC 79956

QY 75 ArgMetLeuPheMetAlaProProGlnProGlnProSerValAspSerValProThrPro 94
Db 79955 CGCATGCTGTTTCATGGCCCGCGAGGCGCGCGCGCGCGG---AGGGCTCCGGTTGTA 79899

QY 95 ValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaGlyLeuSer 114
Db 79898 GCGAGCGGCG---GGTTCGGGAACGAGCGCGCGCATCGCGCGCTGCGCTCTCG 79842

QY 115 ArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlnArgGlySerHis 134
Db 79841 CGTTCCTCCAGCACGGCGTCTCGTGGAC-----GCGGCGGCTCAGGACGAGGCTC 79788

QY 135 ThrCysArgSer----- 138
Db 79787 GCGTCCGCTCCGACGTCGCGGAGGTGAGGAATCGAAACCTAGAAATTCGAATTCGAAA 79728

QY 139 -----Tyr-AspAlaAl 142
Db 79727 AATGATCCAGATCTCTAATCACCGGTGTCTGAAGCGTGGTGTGTTTGTGATGTTCT 79668

QY 142 aGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGlySerG1 162
Db 79667 ACAGGAGGCGAGCGAGCAT-----GTCCCGGAGGCTCCCGGAGGACTCGCGGAGCGG 79614

QY 162 yProAspArgGlu----- 166
Db 79613 CCGTGACCGGAGAGGTGGGAGAAATTCCTCCTCATTTCGGGCAAAATTTCTGAAATTATG 79554

QY 167 -----Ar 167
Db 79553 CGTTCGACATGCCCTGACGCCGAATCACCCCTCGATTTTTCAAATTTTCAATTCAG 79494

QY 167 gArgGluThrThrProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAspLe 187
Db 79493 GAGAGAAACAACTCCATCAAGCTTTCTCCCGGAGAGGTGAGCGATCTGGAGTCTGGATCT 79434

QY 187 uAla---GlyHisLysThrGlyProSerLeuPro---AlaAlaThrProAlaAlaGluLe 205
Db 79433 GGCTGAGGACAGAGCGAGCGCTCCACTACTCTTCTCGGCAACAGCTCAGCACAGCA 79374

QY 205 uIle-----ValProAlaHisGluIleGlnGluPhePheAlaAlaG1 221
Db 79373 AGCCACGCGCGGAAGATTCCGCGCGCGGAGATCGAGGCGTTCTTCGCGGCGGCGCA 79314

QY 221 uAlaAlaGlnAlaLysArgPheAlaSerLys----- 231
Db 79313 GGAGGCTGAGGCCAAGCGCTTCGCCGCCAAGTAAGTATATACAGAAATTTTACCCTCGC 79254

QY 231 ----- 231
Db 79253 TGTAACATCTAGTAAAGTTACTGCGGCCCTAGCTAGCTTCTTCTTCTCGATGC 79194

QY 231 ----- 231
Db 79193 TCGCGGTGGCGTGAATTTGCGCTGCTGCCGGGAGCAAGATAATAACGTGCGTGC 79134

QY 232 -----TyrAsnPheAspPheValArgGlyValProLeuAspAlaG1 245
Db 79133 TTCTTCCCATGGCCGCGAGTACAACCTTCGACGTCGTTTCGCGGCGTGCCCTCGACGCC-- 79076

QY 245 yGlyArgPheGluTrpAlaProValValSer 255
Db 79075 -GGTCGGTTCGAGTGGACTCCGGTGGTCAGC 79046

RESULT 4
AX406682 AX406682 568 bp DNA linear PAT 14-JUN-2002
LOCUS
DEFINITION Sequence 5 from Patent WO0228893.
ACCESSION AX406682
VERSION AX406682.1 GI:21439630
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1
AUTHORS Droual, A.-M., Frankard, V.M.-N., Inze, D., Mironov, V., Peres, B.A.D.
and Hatzfeld, Y.
TITLE Plant cyclin-dependent kinase inhibitors
JOURNAL Patent: WO 0228893-A 5 11-APR-2002;
CROPDESIGN N V (BE)
FEATURES
source
1. 568
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Alignment Scores:
Pred. No.: 2.54e-10 Length: 568
Score: 399.50 Matches: 95
Percent Similarity: 81.90% Conservative: 0
Best Local Similarity: 81.90% Mismatches: 14
Query Match: 30.64% Indels: 7
DB: 6 Gaps: 7

US-09-993-808B-2 (1-256) x AX406682 (1-568)

QY 146 GlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGlySerGlyProAspArg 165
Db 8 GGGGATCACGTC---CCGGATGTCGTCNCGCGAGCAACTCGGGAGCGTCCCGGACCGC 64

QY 166 GluArgArgGluThrThrPro---SerSerArgAlaHis---GlyGluLeuSerAspLeu 183
Db 65 GAGAGGAGAGACGACGCCATCTGTCGAGCGCGGCGCACCGCGCGGAGCTCAGCGATCTG 124

QY 184 GluSerAspLeuAlaGly---HisLysThrGly---ProSerLeuProAla---AlaThr 200
Db 125 GAGTCGATCTGGTGGGGCGGCGAGAGACTGGTCTGCTCGTCGCGCGGAGCAACA 184

QY 201 ProAlaAlaGluLeuIleValProProAlaHisGluIleGlnGluPhePheAlaAla 220
Db 185 TCGGCTGCGGAGCTGATCTGTCGCCGCCAGCACAGAGATCCAGGAATCTTCGCGGCGGCC 244

QY 221 GluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyVal 240

```
Db 245 GAGCGGCGCCATGCCAAACGCTTTGCTTCCAAAGTACAACTTCGACTTCGTCCGCGCGGTG 304
QY 241 ProLeuAspAlaGlyGlyArgPheGluTrpAlaProValSerIle 256
    |||||
Db 305 CCCCTCGACGCC---GGCCGGTTCGAGTGGACGCCAGGGGTGAGCATC 349

RESULT 5
AX406731
LOCUS AX406731 681 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 54 from Patent WO0228893.
ACCESSION AX406731
VERSION AX406731.1 GI:21439657
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE
AUTHORS Droual,A.-M., Frankard,V.M.-N., Inze,D., Mironov,V., Peres,B.A.D.
and Hatzfeld,Y.
TITLE Plant cyclin-dependent kinase inhibitors
JOURNAL Patent: WO 0228893-A 54 11-APR-2002;
CROPDESIGN N V (BE)
FEATURES
    source
        1..681
        /organism="synthetic construct"
        /mol_type="unassigned DNA"
        /db_xref="taxon:32630"
        /note="probe or primer"

ORIGIN
Alignment Scores:
Pred. No.: 7.28e-05 Length: 681
Score: 279.00 Matches: 98
Percent Similarity: 43.51% Conservative: 26
Best Local Similarity: 34.39% Mismatches: 71
Query Match: 21.40% Indels: 90
DB: 6 Gaps: 13

US-09-993-808B-2 (1-256) x AX406731 (1-681)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
    |||||
Db 1 ATGGGAAG---AAGAAGAAGCGCGACGGCGCGCGGCGGAGGAGCGCGGGTGGT 57

QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyVal 40
    |||||
Db 58 GTCGGC-----GGCGTCCGTACGCGG----- 78

QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaAlaValSer 60
    |||||
Db 79 GCGCGCGTCACGGCGGAGGAGGTGGTGGCGAGCGCGGAGGAGGTGTGTTGGTGGGC 138

QY 61 AlaGlyGlyAspGlyGlySer-----CysTyrIleHisLeu 72
    |||||
Db 139 CGTGGCGGTGGCGGTGGCAGTGGCGGAGACGATGGCGAGGCGGATGCTATCTGCGTCTG 198

QY 73 ArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAspSerValPro 92
    |||||
Db 199 CGGAGCAGGAGGCTG----- 213

QY 93 ThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGly 112
    |||||
Db 214 ---CCCTTCGTGGCGCGCGGTGGTGTGTCGTCGCGGAGGAGGAGCGCTCGGTGATTCG 270

QY 113 LeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlnArgGly 132
    |||||
Db 271 GTGGCGGAGGCGGCTTCGTCGTCGTCGTCGCGGCGGTGGAATTGTTGGGC----- 321

QY 133 SerHisThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuValAsp 152
    |||||
Db 322 -----TGT-----TCTGGTGAGGAGGAGGCTATGCGCCGAG 351

QY 153 ValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgArgGluThrThrPro 172
```

```
Db 352 AAGNNGAGCGCG-----ACGACGCGG 372
    |||||
QY 173 SerSerArg-----AlaHisGlyGluLeuSerAspLeuGluSerAsp----- 186
    |||||
Db 373 TCGAGCGCGCGCGCGCGGAGACGCGGACTCGAGCGACGCGGAGTCAAACAGGAGGCC 432
    |||||
QY 187 -----LeuAlaGlyHisLysThrGlyPro 194
    |||||
Db 433 AAGCAGCAATGTCCCGCGGAGTTCGACGACCTCAGCAGTGCATTCACGCGGGA--- 489
    |||||
QY 195 SerLeuProAlaAlaThrProAlaAlaGluLeuIleValProPro-----AlaHisGlu 212
    |||||
Db 490 -----GCGACGACGAGGAGCTTCAGGATGATGGCACCGCGCGCGCGGCGGCGAGAG 540
    |||||
QY 213 IleGlnGluPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAlaSerLysTyr 232
    |||||
Db 541 ATCGAGGAGTCTCGCCGCTCGGAGAGGTTCGAGGCGCGAGCGCTTCGCGCCCAAGTAC 600
    |||||
QY 233 AsnPheAspPheValArgGlyValProLeuAspAlaGly-----GlyArgPheGluTrp 250
    |||||
Db 601 AACTTCGACGTGGTGGCGCGCTGCGCTCGACGCGCGCGCGCGCGGCTTCGAATGG 660
    |||||
QY 251 AlaProValValSer 255
    |||||
Db 661 ACCGCGGTGGGCAGC 675
```

```
RESULT 6
AX406733
LOCUS AX406733 1073 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 56 from Patent WO0228893.
ACCESSION AX406733
VERSION AX406733.1 GI:21439658
KEYWORDS Oryza sativa
SOURCE Oryza sativa
ORGANISM Oryza sativa
REFERENCE
AUTHORS Droual,A.-M., Frankard,V.M.-N., Inze,D., Mironov,V., Peres,B.A.D.
and Hatzfeld,Y.
TITLE Plant cyclin-dependent kinase inhibitors
JOURNAL Patent: WO 0228893-A 56 11-APR-2002;
CROPDESIGN N V (BE)
FEATURES
    Location/Qualifiers
        1..1073
        /organism="Oryza sativa"
        /mol_type="unassigned DNA"
        /db_xref="taxon:4530"

ORIGIN
Alignment Scores:
Pred. No.: 0.00493 Length: 1073
Score: 242.00 Matches: 117
Percent Similarity: 38.20% Conservative: 27
Best Local Similarity: 31.03% Mismatches: 90
Query Match: 18.56% Indels: 144
DB: 6 Gaps: 19

US-09-993-808B-2 (1-256) x AX406733 (1-1073)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
    |||||
Db 1 ATGGGAAG---AAGAAGAAGCGCGACGGCGCGCGGCGGAGGAGCGCGGGTGGT 57

QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyVal 40
    |||||
Db 58 GTCGGC-----GGCGTCCGTACGCGG----- 78

QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaAlaValSer 60
    |||||
Db 79 GCCCGCGTCACGGCGGAGGAGGTGGTGGCGAGCGCGGAGGAGGTGTGTTGGTGGGC 138
```



```
US-09-993-808B-2 (1-256) x AK064723 (1-1045)

QY      1 MetGlyLysTyrMetArgLysCysArgGlyAlaAAGlyAlaGluValAlaAAlaValGlu 20
Db      88 ATGGGGAAGTACATCGGGAAG-----GGGAAGGTGTGCGGGGAGGTGCGGTGATGGAG 141

QY      21 Val---ThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGly 39
Db      142 GTGGGGGGGGCGTCTCGCGCTCCGACCCGCTCCGACACGCTCGCGCTGCAGCGGACG 201

QY      40 ValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaVal 59
Db      202 ACCTCGTCGCAGAGCCGCGGAGAG-----GGGAGGGGGGACCCCGGTGCGGGCGCG 255

QY      60 SerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeupheMet 79
Db      256 GCGCGGGGGCGGAG-----TACCTCGAGCTCAGGAGCGCGAGGCTCGAGAAG 303

QY      80 AlaProProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAsp 99
Db      304 CCGCCTCCGCACACGCCGCG-----GCGAAGGAG 333

QY      100 GlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThr 119
Db      334 AAGGAGACCGCCAGGAGGGCTTCGCGCGCCGCGCCGCGCGCGGTGAGGAT-GCCGCGCGC 392

QY      120 AlaSerSerValAsnLeuGlyLeuGlyGlyGlnArgGly-SerHisThrCysArgSerTy 139
Db      393 GCCGCAAGCGCGGAGGAGTTCGAGCGGAGGTCGAGGTGTCCTTC----- 438

QY      139 rAspAlaAlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSe 159
Db      439 -----GGCGACAACGTTCTT-----GACCTCGA 461

QY      159 rGlySerGlyProAspArgGluArgGluThrThrProSerSer-----ArgAlaHi 177
Db      462 CGGCGACGCCATGGAGAGGAGTAGCAGGGAGACAACGCCCTTCAGTTTAATTAGGAGCTC 521

QY      177 sGlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuPr 197
Db      522 AGAAATGATAAGCACCCCTGGCTCCACAAC-----AAAACCAACACCTCGATCAG 572

QY      197 oAlaAlaThrProAlaAlaGluLeuIle-----ValProProAlaHisGluIleGl 214
Db      573 TTCCCGGCGCAGAAATGGAGACCTCTGTTTGTGCTTACGTTCCGAGTTCTCTTGAGATGGA 632

QY      214 nGluPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPh 234
Db      633 AGAGTTCTTTGAGCTGCTGAACAACAGCAACATCAGGCTTTTCAGAGAGAGGTATAACTT 692

QY      234 eAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGluTrp----- 250
Db      693 CTGTCTCTGTGAACGACTGCCCACTTCCT-----GGACGGTACGAATGGACAAGGCTAGA 746

QY      251 -----AlaProValValSer 255
Db      747 CTGCTAGATTTCATCTTGAGAGCTCCATTGATCTCT 783

RESULT 10
AK073804
LOCUS
DEFINITION
  Oryza sativa (japonica cultivar-group) cDNA clone: J033067M05, full
  insert sequence.
ACCESSION
  AK073804
VERSION
  AK073804.1 GI:32983827
KEYWORDS
  FLI_CDNA; CAP trapper.
SOURCE
  Oryza sativa (japonica cultivar-group)
  Oryza sativa (japonica cultivar-group)
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
  Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
  1
```

AUTHORS

The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-Length cDNA Project Team.,
Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group: Ootomo,Y., Murakami,K.,
Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN.,
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Yoshino,M. and Hayashizaki,Y.

TITLE

Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)
22752273
12869764

REFERENCE

2 (bases 1 to 1027)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
Kodama,T., Kojima,K., Kurohara,C., Kusumegi,T., Li,C., Lu,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miura,J., Miyazaki,A.,
Masuda,H., Matsubara,K., Matsuyama,T., Mizuno,K., Miyazaki,A.,
Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K.,
Numasaki,R., Ohneda,E., Ohtsuki,K., Oka,M., Ooka,H.,
Osato,N., Ota,Y., Otomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K.,
Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,
Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
Sugiyama,A., Suzuki,K., Tagami,M., Tagami-Takeda,Y.,
Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,
Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
Yoshimura,A.

Direct Submission

JOURNAL

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT

This clone is one of the 28K full-length cDNA clones from japonica
rice.
URL : <http://cdna01.dna.affrc.go.jp/cdna/>
NIAS Rice Full-Length cDNA Project Team: Kikuchi,S., Satoh,K.,
Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J.,
Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T.,
Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and
Yamamoto,M.
FAIS Genome Sequencing & Analysis Group: Ootomo,Y., Iida,Y.,
Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M.,
Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Masuda,H., Miura,J.,
Mizuno,K., Narikawa,R., Niikura,J., Oka,M., Ryu,R., Sugano,S.,
Sugiyama,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S.,
Yoshimura,A., Matsubara,K. and Murakami,K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K.,
Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T.,
Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K.,
Hiraoka,T., Hori,F., Iida,J., Imamura,K., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J.,
Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Osato,N.,

Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.
Location/Qualifiers
1. .1027
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J033067M05"

FEATURES
source
ORIGIN
Alignment Scores:
Pred. No.: 0.00716 Length: 1027
Score: 238.00 Matches: 82
Percent Similarity: 40.00% Conservative: 22
Best Local Similarity: 31.54% Mismatches: 76
Query Match: 18.25% Indels: 80
DB: 8 Gaps: 11

US-09-993-808B-2 (1-256) x AK073804 (1-1027)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
Db 130 ATGGCAAGTACATGCGCAAGCCAAAGTGTCTCCGCGAGGTGGTGGCCGCC 189

QY 21 ValThrGlnVal-----ValGlyValArgThrArgSerArgSerAlaAlaAla 36
Db 190 GTCATGGAGCTCGCCGGCGCGCTCGGGTGGCGACCCCGCGCCCTCCCTCGCGCTG 249

QY 37 ThrGlyGlyValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAla 56
Db 250 -----CAGAAGAGG----- 258

QY 57 AlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMet 76
Db 259 -----CAGGGCGGGGAG-----TACCTCGAGCTCAGGAGCCGCGAGG 294

QY 77 LeuPheMetAlaProGlnProGlnProSerValAspSerValProThrProValGlu 96
Db 295 CTCGAGAAGCTCCCTCTCTCCCGCGCG-----CCGCGCGCGAGGAGG 336

QY 97 AlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCys 116
Db 337 AGGGCGACGGTGGCTGCGACTGCTGATGCGACGCGCGCG----- 378

QY 117 SerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlnArgGlySerHisThrCys 136
Db 378 ----- 378

QY 137 ArgSerTyrAspAlaAlaGluAla-----GlyGlyAspHisValLeuValAsp 152
Db 379 -----GAGAGCGCGAGGCGAGGTGCTCGTTCGGGGGGGAGAACGTCCTC---GAG 426

QY 153 ValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThrPro 172
Db 427 CTGAGGCGCATG-----GAAAGGAATACCAGGAGACGACACCT 465

QY 173 SerSer-----ArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHis 190
Db 466 TGCAGCTTGATCAGGAGACCCCGATACGATAGCACCCCTGGATCTACCAAGGCGCAGC 525

QY 191 LysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeuIleValProProAla 210
Db 526 CACTCGAGTTCTCATTTCAAGGTGCAACACCCGTCGCCACACATATTATCCAGCATCA 585

QY 211 HisGluIleGlnGluPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAlaSer 230
Db 586 GCAGAGCTGGAGCGTTCTTCGCTGCCGAGAGCAACGCGCAACGACGAGGTTTCATCGAC 645

QY 231 LysTyrAsnPheAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGluTrp 250

Db 646 AAGTATAACTTTGATCCTGTGAATGACTGCTCTTCCC-----GGCCGGTTGAATGG 699

RESULT 11
AX406720
LOCUS AX406720 1067 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 43 from Patent WO0228893.
ACCESSION AX406720
VERSION AX406720.1 GI:21439647
KEYWORDS Oryza sativa
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1
REFERENCE Droual, A.-M., Frankard, V.M.-N., Inze, D., Mironov, V., Peres, B.A.D.
AUTHORS and Hatzfeld, Y.
TITLE Plant cyclin-dependent kinase inhibitors
JOURNAL Patent: WO 0228893-A 43 11-APR-2002;
CROPDESIGN N V (BE)
FEATURES Location/Qualifiers
source
1. .1067
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"

ORIGIN
Alignment Scores:
Pred. No.: 0.0101 Length: 1067
Score: 235.00 Matches: 82
Percent Similarity: 41.15% Conservative: 25
Best Local Similarity: 31.54% Mismatches: 73
Query Match: 18.02% Indels: 80
DB: 6 Gaps: 12

US-09-993-808B-2 (1-256) x AX406720 (1-1067)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
Db 39 ATGGCAAGTACATGCGCAAGCCAAAGTGTGTCTCCGCGAGGTGGTGGCCGCCGCC 98

QY 21 ValThrGlnVal-----ValGlyValArgThrArgSerArgSerAlaAlaAla 36
Db 99 GTCATGGAGCTCGCCGGCGCGCTCGGGTGGCGACCCCGCGCGCTCCCTCGCGCTG 158

QY 37 ThrGlyGlyValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAla 56
Db 159 -----CAGAAGAGG----- 167

QY 57 AlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMet 76
Db 168 -----CAGGGCGGGGAG-----TACCTCGAGCTCAGGAGCCGCGAGG 203

QY 77 LeuPheMetAlaProGlnProGlnProSerValAspSerValProThrProValGlu 96
Db 204 CTCGAGAAGCTCCCTCTCTCCCGCGCG-----CCGCGCGCGAGGAGG 245

QY 97 AlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCys 116
Db 246 AGGGCGACGGTGGCT-----GCGACTGCT 272

QY 117 SerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlnArgGlySerHisThrCys 136
Db 273 GATGCGACGCGAGC----- 287

QY 137 ArgSerTyrAspAlaAlaGluAla-----GlyGlyAspHisValLeuValAsp 152
Db 288 -----GAGAGCGCGAGGCGAGGTGCTCGTTCGGGGGGGAGAACGTCCTC---GAG 335

QY 153 ValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThrPro 172
Db 336 CTGGAGGCCCATG-----GAAAGGAATACCAGGAGACGACACCT 374

QY 173 SerSer-----ArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHis 190
Db 375 TGCAGCTTGATCAGGACCCCGATACGATTAGCACCCCTGGATCTACCAAGGCGCAGC 434
QY 191 LysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeuValProProAla 210
Db 435 CACTCGAGTTCTCATTTGCAAGGTGCAAAACACCGTGGCCACAACTATTTCAGCATCA 494
QY 211 HisGluIleGlnGluPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAlaSer 230
Db 495 GCAGAGCTGGAGCGTTCTTCGCGCCGAGAGCAACGCGCAACGACAGGCTTTCATCGAC 554
QY 231 LysTyrAsnPheAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGluTyr 250
Db 555 AAGTATAACTTTGATCCTGTGAATGACTGCCCTCTTCCC-----GGCCGRTTTGAATGG 608
RESULT 12
AX463079
LOCUS AX463079 797 bp DNA linear PAT 15-JUL-2002
DEFINITION Sequence 20 from Patent WO0250292.
ACCESSION AX463079
VERSION AX463079.1 GI:21886085
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Fowke,L.C., Wang,H. and Zhou,Y.
TITLE Modulation of plant cyclin-dependent kinase inhibitor activity
JOURNAL Patent: WO 0250292-A 20 27-JUN-2002;
University of Saskatchewan Technologies (CA); Minister of
Agriculture and Agrifood Canada (CA)
FEATURES
source location/Qualifiers
1..797
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
ORIGIN
Alignment Scores:
pred. No.: 0.0578 Length: 797
Score: 215.50 Matches: 79
Percent Similarity: 42.32% Conservative: 34
Best Local Similarity: 29.59% Mismatches: 93
Query Match: 16.53% Indels: 61
DB: 6 Gaps: 12
US-09-993-808B-2 (1-256) x AX463079 (1-797)
QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
Db 10 ATGGGGAATACATGAAGAAATCAAAG-----ATACTGGCGATATCAGCGTCATGGAA 63
QY 21 ValThrGlnValVal-----GlyValArgThrArgSerArgSerAlaAlaAla 36
Db 64 GTCTCTAAAGCAACAGCTCCAAGTCCAGGTGTTTCGAACCAAGCGCGCTAAACCCCTAGCC 123
QY 37 ThrGlyGlyValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAla 56
Db 124 TTG-----AAGCGGCTT-----AATTCCTCC 144
QY 57 AlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMet 76
Db 145 GCCGCTGATTTCAGCTCTACTTAACGACTCTTCTTGTGTATCTTCAGCTCCGTAGCCCGCT 204
QY 77 Leu-----PheMetAlaProProGlnProGlnProSerValAspSerVal 91
Db 205 CTCGAGAAACCCCTCTTCGCTGATTGAACCGAAACAGCGG---CCGAGAGTTTAC----- 255
QY 92 ProThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAla 111

Db 256 -----AGATCGGAATTAAAGACTCTGGT 279
QY 112 GlyLeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlnArg 131
Db 280 TCCAGGTCTCGCGTTGACTCGGTTAACTCG----- 309
QY 132 GlySerHisThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuVal 151
Db 310 ---GCTCCTGTAGCTCAGAGCTCTAATGAAGATGAATGTTTGAACAATTTCGTGAGTGTC 366
QY 152 AspValSerAlaAlaSerAsnSer---GlySerGlyProAspArgGluArgGluThr 170
Db 367 CAAGTTTCTTGTGGTGAAACACAGTCTCGGTTTGAATCAAGACACAGCACAAAGGAGAGC 426
QY 171 ThrProSerArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHis 190
Db 427 ACGCCTTGTAAAC-----TTTGTGAGGATATGGAGATCATGGTTACACACAGG 474
QY 191 LysThrGlyProSerLeuProAlaAlaThrPro-----AlaAlaGluLeuIle 206
Db 475 TCTAGCACGAGGTCGATGTGCAGAGCAACCAAGAGGTACACAAGGGAACAAGATAACGTG 534
QY 207 ValProProAlaHisGluIleGlnGluPhePheAlaAlaGluAlaAlaGlnAlaLys 226
Db 535 ATCCCGACCACTAGTAGAATGGAGGAGTTCTTTGCATATGCAGAGCAGCAGCAACAGAGG 594
QY 227 ArgPheAlaSerLysTyrAsnPheAspPheValArgGlyValProLeuAspAlaGlyGly 246
Db 595 CTATTTCATGGAGAAAGTACAACCTTCGACATTGTGAATGATATCCCGCTC-----AGCGGA 648
QY 247 ArgPheGluTyrAlaProVal 253
Db 649 CGTTACGAATGGGTGCAAGTC 669
RESULT 13
ATH301554 669 bp mRNA linear PLN 26-JUL-2001
LOCUS Arabidopsis thaliana mRNA for cyclin-dependent kinase inhibitor 3
DEFINITION (krp3 gene).
ACCESSION AJ301554
VERSION AJ301554.1 GI:14422288
KEYWORDS cyclin-dependent kinase inhibitor; krp3 gene.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS De Veylder,L., Beeckman,T., Beeckman,T., Krols,L., Terras,F.,
Landrieu,I., van der Schueren,E., Maes,S., Naudts,M. and Inze,D.
TITLE Functional analysis of cyclin-dependent kinase inhibitors of
Arabidopsis
JOURNAL Plant Cell 13 (7), 1653-1668 (2001)
MEDLINE 21342510
PUBMED 11449057
REFERENCE 2 (bases 1 to 669)
AUTHORS de Veylder,L.
TITLE Direct Submission
JOURNAL Submitted (31-OCT-2000) De Veylder L., Departement Plantegenetica,
Vlaams Interuniversitair Instituut voor Biologie (VIB), K.L.
Ledeganckstraat 35, B-9000 Gent, BELGIUM
FEATURES
source location/Qualifiers
1..669
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
gene 1..669
/gene="krp3"
CDS 1..669
/gene="krp3"
/function="cell cycle control"
/codon_start=1

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
Db 9 ATGGGAAATACCTGAGAAAAGT-----GGCCATGTTATG 44
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyVal 40
Db 45 GAGTCTTCTCTTGGTGTCTGTAAGTACGACAACTTAGCA----- 89
QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaAlaValSer 60
Db 90 -----CTTCAACGCTCTTCACTTCTTCTTCAACTCCACCACCTCACTGCC 140
QY 61 AlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAla 80
Db 141 TCTGCTTCTGAT-----TCTTGTACCTTCAACTTCTTCCCGCCCTTCAC---AAA 191
QY 81 ProProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAspGly 100
Db 192 CCACCCACCCCGATCCCTGCCCTAACTCACACCCAC----- 230
QY 101 AlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThrAla 120
Db 231 -----TCTGCTTCTGTT 242
QY 121 SerSerValAsnLeuGlyLeuGlyGlyGlnArgGlySerHisThrCysArgSerTyrAsp 140
Db 243 GATGAATCTCATTTCTGACAATAATCTTCATTTTCAACACACTCATAGAAGC----- 296
QY 141 AlaAlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGly 160
Db 297 ACTAGAGAAGCACACCTTGCAGTCTGCTGAGGAGGTGGACGAATGGTCAACCTTGGT 356
QY 161 SerGlyProAspArgGluArgArgGluThrThrProSerSerArgAlaHisGlyGluLeu 180
Db 357 TCTGCGAAGGCAACAGAGATTAAACACAACTCAAGAGAGGAAGTATATCCTT 416
QY 181 SerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThr 200
Db 417 AGAAT----- 422
QY 201 ProAlaAlaGluLeuIleValProProAlaHisGluIleGlnGluPhePheAlaAla 220
Db 423 -----ATCCCTTCAGCTCATGAATCGAGGACTTTTACCTTTGCT 464
QY 221 GluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyVal 240
Db 465 GAGCAGCAACAACACGCCTTTTATGGACAAGTACAACCTTTGATGTGTAATGATGA 524
QY 241 ProLeuAspAlaGlyGlyArgPheGluTrpAlaProVal 253
Db 525 CCGCTT-----TCTGGCCGTTATGAATGATGATTAGGTTG 557

RESULT 15
AX008805
LOCUS AX008805 804 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 15 from Patent WO9964599.
ACCESSION AX008805
VERSION AX008805.1 GI:9996264
KEYWORDS
SOURCE Chenopodium rubrum (red goosefoot)
ORGANISM Chenopodium rubrum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Amaranthaceae; Chenopodium.

REFERENCE 1
AUTHORS Fowke, L.C., Wang, H. and Crosby, W.L.
TITLE Cyclin-dependent kinase inhibitors as plant growth regulators
JOURNAL Patent: WO 9964599-A 15 16-DEC-1999;
FOWKE LARRY C (CA); WANG HONG (CA); AGRICULTURE AND AGRIFOOD CANAD
(CA); CANADA NAT RES COUNCIL (CA); CROSBY WILLIAM L (CA); UNIV
SASKATCHEWAN TECHNOLOGIES (CA)
Location/Qualifiers

FEATURES

source 1.804
/organism="Chenopodium rubrum"
/mol_type="unassigned DNA"
/db_xref="taxon:3560"
ORIGIN
Alignment Scores:
Pred. No.: 0.302 Length: 804
Score: 199.50 Matches: 67
Percent Similarity: 42.24% Conservative: 31
Best Local Similarity: 28.88% Mismatches: 108
Query Match: 15.30% Indels: 26
DB: 6 Gaps: 7

US-09-993-808B-2 (1-256) x AX008805 (1-804)
QY 24 ValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGly-ValAlaLysVa 43
Db 18 GTGGTAGGAGTTAAACACAGAGCTCGAGACTGCCCTAGCTATGGCGGAGCTGCT----- 72
QY 43 lAlaProArgLysArgAlaProAlaGlyGluProAlaAlaAlaValSerAlaGlyGl 63
Db 73 -----ACTCCAACTTCGTCCTCCGGCGAAGAGATCAAGAGGTTTC 113
QY 63 yAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAlaProGl 83
Db 114 GAAGTCGTCGTATAATATCTCTCAACTAAGAGTCTCGAAAGAAATTTGTCGGCGCGGA 173
QY 83 nProGlnProSerValAspSerValProThrProValGluAlaAlaAspGlyAlaAlaGl 103
Db 174 GAATTTCCCGGAATTAGAAACG-----ACGCCGTTGGAAGTTGCGGCGGTTGTTGAGGA 227
QY 103 yGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThrAlaSerSerVa 123
Db 228 GGAAGAG-----GTTGCGAATTGCTCGAGT-----AGCGAGGT 260
QY 123 lAsnLeuGlyLeuGlyGlyGlnArgGlySerHisThrCysArgSer---TyrAspAlaAl 142
Db 261 AATTACTACAGCTAGTCCGATTTTCCGCCGCTCTTGTGTCACGAATATGATCAGTT 320
QY 142 aGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGlySerGl 162
Db 321 GAGTTCTAGCGAGCCAGAGTAGTTAAGGATGATGATGTTGGGAAATCGTACAGCAGA 380
QY 162 yProAspArgGluArgArgGluThrThrProSerSerArgAlaHisGlyGluLeuSerAs 182
Db 381 TCCAGAGTTGAGAGTGGTGGCGCTCGTCAAAGCAAAAG-----GAGAGCCATAG 431
QY 182 pLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThrProAl 202
Db 432 AACAGAAGCGAGAGAGAGCTACAAAATTAGACGACCGAGGATTATCCGCGCAGAAATCAAC 491
QY 202 aAlaGluLeuIleValProProAlaHisGluIleGlnGluPhePheAlaAlaGluAl 222
Db 492 GGTACAGATCAAGATGCCGCTGATTTCAGAAATCGAAGAAATCTTTGCTGTTGCTGAAA 551
QY 222 aAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyValProLe 242
Db 552 AGATCTCCAGAAACGCTTCAGCGAAAGTACAATTTTCGACATAGTTAAGGACGCTGCCACT 611
QY 242 uAspAlaGlyGlyArgPheGluTrpAlaProVal 253
Db 612 GAAA-----GGTCGTTATGATGGGTTCCAATA 639

Search completed: October 2, 2004, 11:11:36
Job time : 3207 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2004, 08:35:39 ; Search time 366 Seconds
(without alignments)
2971.417 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 1304
Sequence: 1 MGKYMRCRGAGAEVAAVE.....VRGVPLDAGGRFEWAPVWSI 256

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09993808/runat_01102004_171701_13019/app_query.fasta_1.455
-DB=N Geneseq 29Jan04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09993808 @CGN_1_1_708 @runat_01102004_171701_13019 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002s.*
7: Geneseqn2003as.*
8: Geneseqn2003bs.*
9: Geneseqn2003cs.*
10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 1304 | 100.0 | 1372 | 7 ABV74603 | Abv74603 Maize CKI |
| 2 | 669 | 51.3 | 1242 | 6 ABK93958 | Abk93958 DNA encod |
| 3 | 664 | 50.9 | 841 | 7 ABV74605 | Abv74605 Maize CKI |
| 4 | 612 | 46.9 | 723 | 3 AAA95281 | Aaa95281 Corn cycl |
| 5 | 413 | 31.7 | 572 | 3 AAN02390 | Aan02390 Corn cycl |
| 6 | 413 | 31.7 | 572 | 3 AAA95276 | Aaa95276 Corn cycl |
| 7 | 411 | 31.5 | 639 | 3 AAA95280 | Aaa95280 Corn cycl |
| 8 | 399.5 | 30.6 | 568 | 6 ABK93954 | Abk93954 Corn zmIC |

| | | | | | | |
|----|-------|------|--------|---|-------------|--------------------|
| 9 | 279 | 21.4 | 681 | 6 | ABK93981 | Abk93981 Full leng |
| 10 | 242 | 18.6 | 1073 | 6 | ABK93982 | Abk93982 Rice OsIC |
| 11 | 235 | 18.0 | 1067 | 6 | ABK93971 | Abk93971 Full leng |
| 12 | 219 | 16.8 | 1169 | 3 | AAA95292 | Aaa95292 Wheat cyc |
| 13 | 215.5 | 16.5 | 797 | 6 | AAD40771 | Aad40771 Arabidops |
| 14 | 204 | 15.6 | 1005 | 3 | AAA95283 | Aaa95283 Corn cycl |
| 15 | 200 | 15.3 | 1089 | 7 | ABV74604 | Abv74604 Maize CKI |
| 16 | 199.5 | 15.3 | 804 | 3 | AAZ29420 | Aaz29420 Chenopodi |
| 17 | 199.5 | 15.3 | 804 | 6 | AAD40768 | Aad40768 Chenopodi |
| 18 | 190.5 | 14.6 | 875 | 2 | AAZ25016 | Aaz25016 Arabidops |
| 19 | 178 | 13.7 | 870 | 3 | AAA95293 | Aaa95293 Arabidops |
| 20 | 178 | 13.7 | 870 | 3 | AAC46465 | Aac46465 Arabidops |
| 21 | 178 | 13.7 | 1290 | 6 | AAD40769 | Aad40769 Arabidops |
| 22 | 163 | 12.5 | 1116 | 3 | AAN02402 | Aan02402 Cyclin de |
| 23 | 163 | 12.5 | 1116 | 3 | AAA95288 | Aaa95288 Soybean c |
| 24 | 152.5 | 11.7 | 110000 | 4 | AAI99683_39 | Continuation (40 o |
| 25 | 151.5 | 11.6 | 493 | 6 | ABK93950 | Abk93950 Rice OsIC |
| 26 | 149 | 11.4 | 984 | 6 | ABZ12222 | Abz12222 Arabidops |
| 27 | 145.5 | 11.2 | 204 | 3 | AAA95286 | Aaa95286 Rice cycl |
| 28 | 145.5 | 11.2 | 204 | 3 | AAN02400 | Aan02400 Rice cycl |
| 29 | 145 | 11.1 | 533 | 3 | AZ29419 | Aaz29419 Arabidops |
| 30 | 145 | 11.1 | 533 | 6 | AAD40767 | Aad40767 Arabidops |
| 31 | 143.5 | 11.0 | 2328 | 7 | ABX56289 | Abx56289 Human NOV |
| 32 | 143.5 | 11.0 | 2511 | 7 | ACC42663 | Acc42663 Human LP3 |
| 33 | 142.5 | 10.9 | 472 | 3 | AAA95284 | Aaa95284 Corn cycl |
| 34 | 141.5 | 10.9 | 4281 | 7 | ACA25471 | Aca25471 Prokaryot |
| 35 | 141.5 | 10.9 | 110000 | 4 | AAI99682_09 | Continuation (10 o |
| 36 | 141.5 | 10.9 | 110000 | 4 | AAI99683_09 | Continuation (10 o |
| 37 | 140.5 | 10.8 | 609 | 3 | AAA95279 | Aaa95279 Wheat cyc |
| 38 | 140.5 | 10.8 | 110000 | 4 | AAI99682_39 | Continuation (40 o |
| 39 | 138.5 | 10.6 | 3921 | 7 | ACA40351 | Aca40351 Prokaryot |
| 40 | 138.5 | 10.6 | 110000 | 4 | AAI99682_06 | Continuation (7 of |
| 41 | 138.5 | 10.6 | 110000 | 4 | AAI99683_06 | Continuation (7 of |
| 42 | 137 | 10.5 | 1143 | 7 | ADA70803 | Ada70803 Rice Gene |
| 43 | 135.5 | 10.4 | 2337 | 7 | ACA40697 | Aca40697 Prokaryot |
| 44 | 135.5 | 10.4 | 110000 | 4 | AAI99682_29 | Continuation (30 o |
| 45 | 135.5 | 10.4 | 110000 | 4 | AAI99683_29 | Continuation (30 o |

ALIGNMENTS

RESULT 1
ABV74603
ID ABV74603 standard; DNA; 1372 BP.
XX
AC ABV74603;
XX
DT 21-FEB-2003 (first entry)
XX
DE Maize CKI_B coding sequence.
XX
KW Maize; cyclin-dependent kinase inhibitor; CKI; CKI_B; plant; corn;
KW crop yield; root size; plant growth; tassel size; ear size;
KW male sterility; endoreduplication; gene; ds.
OS Zea mays.
XX
FH Key
CDS Location/Qualifiers
134..904
/*tag= a
/product= "CKI_B"
WO200281623-A2.
XX
PD 17-OCT-2002.
XX
PF 06-NOV-2001; 2001WO-US044038.
XX
PR 07-NOV-2000; 2000US-0246349P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
PA (ARIZ-) ARIZONA BOARD OF REGENTS.
XX

PI Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;
XX WPI; 2003-058511/05.
DR P-PSDB; ABB98757.
XX Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
PT involved in cell cycle regulation, and useful for altering cell cycle
PT protein content, cell cycle progression, cell number and composition of
PT plants.
XX Claim 1; Page 64-65; 69pp; English.
PS
XX The present sequence is the coding sequence (I) for maize cyclin-
CC dependent kinase inhibitor (CKI), CKI B. (I) is useful for modulating the
CC activity of cyclin-dependent kinase (CDK) in a plant such as a corn,
CC soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice,
CC barley, oil-seed Brassica and millet. Modulating the activity of CDK,
CC preferably modulating downward is useful for providing differential
CC growth in a plant, especially a positive growth advantage and modulating
CC CDK activity upward is useful for increasing crop yield, root size, plant
CC growth, tassel size and/or ear size. Modulating CDK activity is also
CC useful for conferring male sterility and for improving transformation
CC frequencies by increasing the number of cells in cell division. CDK
CC activity can also be modulated for modulating endoreduplication in the
CC endosperm of corn, sorghum, wheat, rice, barley, and millet, where the
CC promoter used is an endosperm-preferred promoter. The cell numbers are
CC modulated in one or more tissues of a plant, comprising root, seed,
CC tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves,
CC stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm,
CC pericarp, meristematic or leaf cells. (I) is also useful for identifying
CC maize CKI interacting proteins, by adducting the nucleic acid sequence to
CC a second nucleic acid sequence encoding a DNA-binding domain
XX
SQ Sequence 1372 BP; 292 A; 396 C; 442 G; 241 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 2.44e-67 Length: 1372
Score: 1304.00 Matches: 256
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-09-993-808B-2 (1-256) x ABV74603 (1-1372)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
Db 134 ATGGGAAGTACATCGCAAGTGCAGGGCGCCGAGGCGCGGAGGTGCGCCCGTCGAG 193
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal 40
Db 194 GTTACGAGGTCTGTCGGCGTCCGACGAGTCCAGGTCCGCGCGCGGACCGCGGTGTC 253
QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaValSer 60
Db 254 GCGAAGTCCCGCCGAGGAGGAGAGGGCGCGCGGGGAGCCTGTCGCGCGGTGAGC 313
QY 61 AlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAla 80
Db 314 GCTGTGGGACGCGCGGAAGTCTACATCCACCTCGTAGCCGCGGTCTGTTTCATGGCA 373
QY 81 ProProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAspGly 100
Db 374 CCGCCTCAGCCGAGCCGTCGGTTGACTCGGTCCGACCCCGGTGGAGGCTGTGATGGC 433
QY 101 AlaAlaGlyGlnGlnGlyAlaAlaLeuAlaGlyLeuSerArgCysSerSerThrAla 120
Db 434 GCTGCAGGACAGCAGGGCGCGGCGTCTCGCGCGGGCTCTCGCGTTGCTCCAGCAGCGG 493
QY 121 SerSerValAsnLeuGlyLeuGlyGlnArgGlySerHisThrCysArgSerTyrAsp 140
Db 494 TCGTCGGTGAACCTTGGGCTTGGGGGTCTAGCGGGGAGCCACACCTGCGCTCCTACGAC 553

QY 141 AlaAlaGluAlaGlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGly 160
Db 554 GCTGCAGAGGCTGGCGGGATCACGTCTCTGGTGTGTCTCGGGCGGAGCACTCCGGG 613
QY 161 SerGlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyGluLeu 180
Db 614 AGCGGCCAGACCGCGAGAGCGAGAGACGACCGCATCGAGCCGGCGGCGAGCTC 673
QY 181 SerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThr 200
Db 674 AGCGATCTGGAGTCTGGATCTGGCGGGGACAAAGACTGGCCCTCGTACCGGCGGCAACG 733
QY 201 ProAlaAlaGluLeuValProProAlaHisGluLeuGlnGluPheAlaAlaAla 220
Db 734 CCGGCTGCGGAGTGTATCGTCCCGCCAGCACACGAGATCCAGGAGTTCTTCGCCCGCGCC 793
QY 221 GluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyVal 240
Db 794 GAGCGGCCCGAGGCCAAGCGCTTGTCTCCAAGTACAACTTCGACTTCGTCCGCGCGGTG 853
QY 241 ProLeuAspAlaGlyGlyArgPheGluTrpAlaProValValSerIle 256
Db 854 CCCCTCGACGCCCGCGCGGTTCGAGTGGGCGCGCGGTGGTTCAGCATC 901
RESULT 2
ABK93958
ID ABK93958 standard; DNA; 1242 BP.
XX
AC ABK93958;
XX 27-AUG-2002 (first entry)
XX DNA encoding OsICK2 protein.
DE
XX Plant; ds; gene; inhibitor of cyclin dependent kinase; ICK.
KW
XX Oryza sativa.
OS
XX WO200228893-A2.
PN
XX 11-APR-2002.
PD
XX 29-JUN-2001; 2001WO-IB001492.
PF
XX 14-JUL-2000; 2000US-0218471P.
PR
XX 13-OCT-2000; 2000US-0241219P.
XX (CROP-) CROPDESIGN NV.
PA
PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
PI Hatzfeld Y;
XX
DR WPI; 2002-471311/50.
DR P-PSDB; ABG65670.
XX
PT Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
PT to screen substrates, drugs or compounds which modulate ICK activity and
PT treat disorders characterized by an insufficient or excessive production
PT of ICK inhibitors.
XX
PS Claim 46; Fig 1; 141pp; English.
XX This invention relates to the DNA and protein sequences of novel isolated
CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
CC the invention may be used for treating disorders characterised by
CC insufficient or excessive production of an ICK inhibitor. The protein of
CC the invention may also be used to screen for naturally-occurring ICK
CC substrates, drugs or compounds which modulate ICK activity, as well as to
CC treat disorders characterised by insufficient or aberrant activity compared to
CC ICK protein, forms which have decreased or aberrant activity compared to
CC ICK wild type protein. The present sequence represents a cDNA molecule
CC encoding an inhibitor of cyclin dependent kinase (ICK) protein of the
CC invention

XX SQ Sequence 1242 BP; 286 A; 319 C; 417 G; 220 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.78e-30 Length: 1242
Score: 669.00 Matches: 161
Percent Similarity: 67.42% Conservative: 19
Best Local Similarity: 60.30% Mismatches: 67
Query Match: 51.30% Indels: 20
DB: 6 Gaps: 10

US-09-993-808B-2 (1-256) x ABX93958 (1-1242)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
Db |||||
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGly--- 39
Db |||||
QY 40 -----ValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGlu 54
Db |||||
QY 138 ACGACGAGGTGCAGGCGCGTCCGCGCGTCCACGAGGAGGAGGAGGCGTGTGCGG 197
Db |||||
QY 55 ProAlaAlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSer 74
Db |||||
QY 198 ACGCGGTGCTGGGACTACTCGCGTACGCGGAGTGTACCTCAGCTGAGGAGC 257
Db |||||
QY 75 ArgMetLeuPheMetAlaProGlnProGlnProSerValAspSerValProThrPro 94
Db |||||
QY 258 CGCATGTGTTTCATGGCCCGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 314
QY 95 ValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSer 114
Db |||||
QY 315 GCGGAGCGGCGG---GGTTCCGGGAACGAGCGGCGGCGCATGCGCGGCTGCGCTCG 371
QY 115 ArgCysSerSerThrAlaSerSerValAsnLeuGlyGlyGlnArgGlySerHis 134
Db |||||
QY 372 CGTGTCTCAGCACGCGCTGCTCGTGGAC-----GCGCGCGCTCAGGACAGGCGCTC 425
QY 135 ThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuValAspValSer 154
Db |||||
QY 426 GCGTGGCGCTCC---GACGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 476
QY 155 AlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThrProSerSer 174
Db |||||
QY 477 TCCGCGAGCGACTCGGCGAGCGGCGTGCAGCGGAGGAGGAGGAGGAGGAGGAGGAGG 536
QY 175 ArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAla---GlyHisLysThrGly 193
Db |||||
QY 537 TTTCTCCCGCGGAGGTGAGCGATCTGGAGTCTGGTGGAGGAGGAGGAGGAGGAGGAGG 596
QY 194 ProSerLeuPro---AlaAlaThrProAlaAlaGluLeuLeu-----ValPro 208
Db |||||
QY 597 CGTCCACTACCTTCTCGGCAACAGCTCAGCACAGCAAGCCAGCGGCGGCGGCGGCGG 656
QY 209 ProAlaHisGluGlnGluPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPhe 228
Db |||||
QY 657 CCGGCGCGCGAGATCGAGCGGTCTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 716
QY 229 AlaSerLysTyrAsnPheAspPheValArgGlyValProLeuAspAlaGlyGlyArgPhe 248
Db |||||
QY 717 GCCGCCAAGTACAACCTTCGACGCTGCTTCGCGGCGGTGCCCTCGACGCGC---GGT 773
QY 249 GluTrpAlaProValValSer 255
Db |||||
QY 774 GAGTGGACTCCGGTGGTTCAGC 794

RESULT 3
ABV74605
ID ABV74605 standard; DNA; 841 BP.
XX

AC ABV74605;
XX
DT 21-FEB-2003 (first entry)
XX
DE Maize CKI_D coding sequence.
XX
KW Maize; cyclin-dependent kinase inhibitor; CKI; CKI_D; plant; corn;
KW crop yield; root size; plant growth; tassel size; ear size;
KW male sterility; endoreduplication; gene; ds.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 159..839
FT /*tag= a
FT /partial
FT /product= "CKI_D"
FT /note= "No stop codon given"
XX
PN WO200281623-A2.
XX
PD 17-OCT-2002.
XX
PF 06-NOV-2001; 2001WO-US044038.
XX
PR 07-NOV-2000; 2000US-0246349P.
XX
PA (PION-) PIONEER HI-BRED INT INC.
PA (ARIZ-) ARIZONA BOARD OF REGENTS.
XX
PI Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;
XX WPI; 2003-058511/05.
DR P-PSDB; ABB98759.
XX
PT Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
PT involved in cell cycle regulation, and useful for altering cell cycle
PT protein content, cell cycle progression, cell number and composition of
PT plants.
XX
PS Claim 1; Page 68-69; 69pp; English.
XX
CC The present sequence is the coding sequence (I) for maize cyclin-
CC dependent kinase inhibitor (CKI), CKI_D. (I) is useful for modulating the
CC activity of cyclin-dependent kinase (CDK) in a plant such as a corn,
CC soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice,
CC barley, oil-seed Brassica and millet. Modulating the activity of CDK,
CC preferably modulating downward is useful for providing differential
CC growth in a plant, especially a positive growth advantage and modulating
CC CDK activity upward is useful for increasing crop yield, root size, plant
CC growth, tassel size and/or ear size. Modulating CDK activity is also
CC useful for conferring male sterility and for improving transformation
CC frequencies by increasing the number of cells in cell division. CDK
CC activity can also be modulated for modulating endoreduplication in the
CC endosperm of corn, sorghum, wheat, rice, barley, and millet, where the
CC promoter used is an endosperm-preferred promoter. The cell numbers are
CC modulated in one or more tissues of a plant, comprising root, seed,
CC tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves,
CC stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm,
CC pericarp, meristematic or leaf cells. (I) is also useful for identifying
CC maize CKI interacting proteins, by adducting the nucleic acid sequence to
CC a second nucleic acid sequence encoding a DNA-binding domain
XX
SQ Sequence 841 BP; 142 A; 247 C; 342 G; 109 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 2.36e-30 Length: 841
Score: 664.00 Matches: 172
Percent Similarity: 75.11% Conservative: 6
Best Local Similarity: 72.57% Mismatches: 31
Query Match: 50.92% Indels: 28
DB: 7 Gaps: 16


```
US-09-993-808B-2 (1-256) x ABV74605 (1-841)
QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaValGlu 20
Db 159 ATGGGAAGTACATGCGCAAGCGCAGGGGGCGCGGAGGGGTGGCCGAGTCGAG 218
QY 21 ValThrGlnValGlyValArgThrArgSerArgSerAlaAlaThr---GlyGly 39
Db 219 GTCTCGCAGGTCGTGCGCGTCCGACGAGGTCAGGTCGCGGCGGCGACCGCGCGGT 278
QY 40 ValAlaLysValAla---ProArgArgLysArgAla-----ProAla----- 52
Db 279 GTCGGAAGGTCGTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 338
QY 53 -----GlyGluProAlaAlaAlaValSerAlaGlyGlyAspGlyGlySer---CysTyr 69
Db 339 ACGTCGGGGAGCGCTGCTGCGGTGGCGCTGGTGGTGGGACGCGGGAAGCTGCTGCTAC 398
QY 70 IleHisLeuArgSerArgMetLeuPheMetAlaProGlnProGlnProSerValAsp 89
Db 399 ATCCACCTGCGGAGCGCGATGCTGCTATGCGAGCACCTCAGCAGCAACCGTCG----- 452
QY 90 SerValProThrProValGluAlaAlaAlaAspGlyAlaAlaGlyGlnGlnGly---AlaAla 108
Db 453 GCGGCTCTGACGCGCGGTGGAGGCTGCT---GGTGGGACACARCAAGGCGGGTGGTGGCG 509
QY 109 LeuAlaAlaGlyLeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGly 128
Db 510 CTCGGGCTGGCCTCTCGCGTGTGCTCCAGCACGCGCTGCTGCTGCTGCTGCTGCTGCT 560
QY 129 GlyGlnArgGlySerHisThrCysArgSerTyrAspAla---AlaGluAlaGlyGlyAsp 147
Db 561 -----GGGGGCCACGCGCTGCGGCTCCGACGCTGCGGCTGCGGAGGTTGACGGGGAT 611
QY 148 HisValLeuValAspValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArg 167
Db 612 CACGTC---CCGATGTGCTACCGCGAGCAACTCGGGGAGCGTCCCGGACCGCGAGAGG 668
QY 168 ArgGluThrThrPro---SerSerArgAlaHis---GlyGluLeuSerAspLeuGluSer 185
Db 669 AGAGAGACGACGCCATCGTCGAGCGGGCGCAGCGGCGGAGCTCAGCGATCTGGAGTCG 728
QY 186 AspLeuAlaGly---HisLysThrGly---ProSerLeuProAla---AlaThrProAla 202
Db 729 GATCTGTGGGGCGGCGAGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 788
QY 203 AlaGluLeuIleValProProAlaHisGluIleGlnGlnPhePheAlaAla 219
Db 789 GCGGAGCTGATGTCGCGCGCCAGCACAGGAGATCCAGGAATTCCTTCGCGGCC 839

RESULT 4
AAA95281
ID AAA95281 standard; cDNA; 723 BP.
AC AAA95281;
XX
DT 17-JAN-2001 (first entry)
XX
DE Corn cyclin-dependent kinase inhibitor coding sequence #3.
XX
KW Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
XX cell growth; herbicide; ss.
OS Zea mays.
XX
FH Key Location/Qualifiers
CDS 3..380
FT /*tag= a
FT /product= "CDKI"
FT /partial
XX
PN W0200060087-A2.
XX
```

```
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US009106.
XX
PR 07-APR-1999; 99US-0128192P.
XX
PA (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI Klein TM, Weng Z, Cahoon RE;
XX
DR WPI; 2000-679375/66.
DR P-PSDB; AAB26250.
XX
PT Cyclin dependent kinase inhibitor sequences, useful for identifying
PT herbicides and plant growth inhibitors.
XX
PS Claim 2; Page 43-44; 58pp; English.
XX
CC The present sequence is the coding sequence for the corn cyclin-dependent
CC kinase inhibitor (CDKI). It was isolated by searching a contig comprising
CC cDNA from corn roots and ear leaf sheath for sequences similar to those
CC encoding the CDKI from Chenopodium rubrum, Caenorhabditis elegans and
CC Arabidopsis thaliana. CDKI is involved in the cell cycle, and may promote
CC or inhibit cell division and growth. The coding sequence and the protein
CC it encodes are useful in the production of transgenic plants which
CC produce increased or decreased amounts of the CDKI protein, in the
CC identification of herbicides, in genetic and physical mapping and in the
CC isolation of the CDKI gene in other organisms
XX
SQ Sequence 723 BP; 160 A; 202 C; 210 G; 151 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.14e-27 Length: 723
Score: 612.00 Matches: 120
Percent Similarity: 96.80% Conservative: 1
Best Local Similarity: 96.00% Mismatches: 4
Query Match: 46.93% Indels: 0
DB: 3

US-09-993-808B-2 (1-256) x AAA95281 (1-723)
QY 132 GlySerHisThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuVal 151
Db 3 GGGAGCCACACCTGCCGCTCCGACGACGCTGACAGGCTGGCGGGGATCAGTCCTGGTG 62
QY 152 AspValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThr 171
Db 63 GTGGATGTCGCGGAGCAACTCCGGAGCGGCCAGCCGAGAGGAGAGAGAGAGAGAG 122
QY 172 ProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHisLys 191
Db 123 CCATCGAGCGGGCGCACCGCGAGCTCAGCGATCTGGAGTCGGATCTGGCGGGGCAAG 182
QY 192 ThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeuIleValProProAlaHis 211
Db 183 ACTGGCCGTCGCTACCGCGCGCAACCGCGGCTGCGGAGCTGATGTCGCCGCCAGCAC 242
QY 212 GluIleGlnGluPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAlaSerLys 231
Db 243 GAGATCCAGGAGTCTTCGCGCGCGCGAGGCGGCCAGGCCAAGCGCTTTGCTTCCAAG 302
QY 232 TyrAsnPheAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGluTrpAla 251
Db 303 TACAACCTCGACTTCGTCCGCGCGGTGCGCCCTCGACCGCGCGCGGTTCGAGTGGCG 362
QY 252 ProValValSerIle 256
Db 363 CCGGTGTCAGCATC 377

RESULT 5
AAN02390
ID AAN02390 standard; DNA; 572 BP.
XX
```

```

121 AGCCGGCGCACGGCGGCGAGCTCAGCGATCTGGAGTCGGATCTGGTGGGCGGAGAG 180
192 ThrGly---ProSerLeuProAlaAlaThrProAla---AlaGluLeuIleValProPro 209
181 ACTGGTGTCTCGTCTCGTCCGCGGACAAACATCGGTTGGGAGCTGATCGTGCCGCCA 240
210 AlaHisGluIleGlnGluPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAla 229
241 GCACAGGAGATCCAGGAATCTTCGGCGCCGCGAGGCGGCCCATGCAACGCTTTGCT 300
230 SerLysTyrAsnPheAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGlu 249
301 TCCAAGTACAACCTCGACTTCGTCCGCGGCGTGCCCCCTCGACGCC---GGCCGGTTCGAG 357

250 TrpAlaProValValSerIle 256
358 TGGACGCCANGGGTCAGCATC 378

RESULT 6
AA95276
ID AAA95276 standard; cDNA; 572 BP.
AC
XX AAA95276;
XX
DT 17-JAN-2001 (first entry)
DE
XX
DE Corn cyclin-dependent kinase inhibitor coding sequence #1.
XX
DE Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
KW cell growth; herbicide; ss.
KW
XX
XX Zea mays.
XX
XX
XX Key Location/Qualifiers
FH 208..366
FT CDS /tag=a
FT /product="CDKI"
FT /partial
XX
XX WO200060087-A2.
XX
XX 12-OCT-2000.
XX
XX 06-APR-2000; 2000WO-US009106.
XX
XX 07-APR-1999; 99US-0128192P.
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
XX Klein TM, Weng Z, Cahoon RE;
XX
XX WPI; 2000-679375/66.
XX P-PSDB; AAB26245.
XX
XX Cyclin dependent kinase inhibitor sequences, useful for identifying
XX herbicides and plant growth inhibitors.
XX
XX Claim 2; Page 39-40; 58pp; English.
XX
XX The present sequence is the coding sequence for the corn cyclin-dependent
XX kinase inhibitor (CDKI). It was isolated by searching a corn silk cDNA
XX library for sequences similar to those encoding the CDKI from Chenopodium
XX rubrum, Caenorhabditis elegans and Arabidopsis thaliana. CDKI is involved
XX in the cell cycle, and may promote or inhibit cell division and growth.
XX The coding sequence and the protein it encodes are useful in the
XX production of transgenic plants which produce increased or decreased
XX amounts of the CDKI protein, in the identification of herbicides, in
XX genetic and physical mapping and in the isolation of the CDKI gene in
XX other organisms
XX
XX Sequence 572 BP; 140 A; 156 C; 176 G; 93 T; 0 U; 7 Other:
SQ

```

Alignment Scores:

Pred. No.: 6.25e-16 Length: 572
Score: 413.00 Matches: 100
Percent Similarity: 79.53% Conservative: 1
Best Local Similarity: 78.74% Mismatches: 18
Query Match: 31.67% Indels: 8
DB: 3 Gaps: 8

US-09-993-808B-2 (1-256) x AAA95276 (1-572)

QY 136 CysArgSerTyrAspAla---AlaGluAlaGlyGlyAspHisValLeuValAspValSer 154
Db 4 TGGCGTCCGACGCTCGCTCGGAGGTTGACGGGGATCACGTC---CCGGATGTCGTC 60
QY 155 AlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrPro---Ser 173
Db 61 ACCGCGAGCAACTCCGGAGCGTCCCGGACCGCGAGAGAGAGACGACGCCATCGTCG 120
QY 174 SerArgAlaHis---GlyGluLeuSerAspLeuGluSerAspLeuAlaGly---HisLys 191
Db 121 AGCCGGGCGCACGGCGGAGCTCAGCGATCTGGAGTCGGATCTGGTGGGGCGGAGAG 180
QY 192 ThrGly---ProSerLeuProAlaAlaThrProAla---AlaGluLeuIleValProPro 209
Db 181 ACTGGTGTCTGCTCGTCCGCGGCGACAAACATCGGTTGCGGAGTGATCGTCCGCCA 240
QY 210 AlaHisGluIleGlnGluPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAla 229
Db 241 GCACAGGAGATCCAGGAATCTTCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 230 SerLysTyrAsnPheAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGlu 249
Db 301 TCCAAGTACAACTTCGACTTCGTCCGCGCGGCGTGCCTCGACGCC---GGCCGGTTCGAG 357
QY 250 TrpAlaProValValSerIle 256
Db 358 TGGACGCCANGGGTCAGCATC 378

RESULT 7

AAA95280
ID AAA95280 standard; cDNA; 639 BP.

XX AC AAA95280;

XX XX 17-JAN-2001 (first entry)

XX DE Corn cyclin-dependent kinase inhibitor coding sequence #2.

XX XX Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
KW cell growth; herbicide; ss.

XX OS Zea mays.

XX FH Key Location/Qualifiers
FT CDS 8..388
FT /*tag= a
FT /product= "CDKI"
FT /partial

XX PN WO2000060087-A2.

XX XX 12-OCT-2000.

XX PF 06-APR-2000; 2000WO-US0009106.

XX PR 07-APR-1999; 99US-0128192P.

XX PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX PI Klein TM, Weng Z, Cahoon RE;

XX DR WPI; 2000-679375/66.

XX DR P-PSDB; AAB26249.

XX

PT Cyclin dependent kinase inhibitor sequences, useful for identifying herbicides and plant growth inhibitors.

XX PS Claim 2; Page 43; 58pp; English.

XX CC The present sequence is the coding sequence for the corn cyclin-dependent kinase inhibitor (CDKI). It was isolated by searching a corn silk cDNA library for sequences similar to those encoding the CDKI from Chenopodium rubrum, Caenorhabditis elegans and Arabidopsis thaliana. CDKI is involved in the cell cycle, and may promote or inhibit cell division and growth.

XX CC The coding sequence and the protein it encodes are useful in the production of transgenic plants which produce increased or decreased amounts of the CDKI protein, in the identification of herbicides, in genetic and physical mapping and in the isolation of the CDKI gene in other organisms

XX SQ Sequence 639 BP; 169 A; 169 C; 192 G; 109 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 9.11e-16 Length: 639
Score: 411.00 Matches: 100
Percent Similarity: 79.53% Conservative: 1
Best Local Similarity: 78.74% Mismatches: 18
Query Match: 31.52% Indels: 8
DB: 3 Gaps: 8

US-09-993-808B-2 (1-256) x AAA95280 (1-639)

QY 136 CysArgSerTyrAspAla---AlaGluAlaGlyGlyAspHisValLeuValAspValSer 154
Db 11 TGGCGTCCGACGCTCGCTCGGAGGTTGACGGGGATCACGTC---CCGGATGTCGTC 67

QY 155 AlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrPro---Ser 173
Db 68 ACCGCGAGCAACTCCGGGAGCGTCCCGGACCGCGAGAGAGAGACGCCATCGTCG 127

QY 174 SerArgAlaHis---GlyGluLeuSerAspLeuGluSerAspLeuAlaGly---HisLys 191
Db 128 AGCCGGGCGCACGGCGGAGCTCAGCGATCTGGAGTCGGATCTGGTGGGGCGGAGAG 187

QY 192 ThrGly---ProSerLeuProAlaAlaThrProAla---AlaGluLeuIleValProPro 209
Db 188 ACTGGTGTCTGCTCGTCCGCGGCGACAAACATCGGTTGCGGAGTGATCGTCCGCCA 247

QY 210 AlaHisGluIleGlnGluPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAla 229
Db 248 GCACAGGAGATCCAGGAATCTTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 307

QY 230 SerLysTyrAsnPheAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGlu 249
Db 308 TCCAAGTACAACTTCGACTTCGTCCGCGGCGTGCCTCGACGCC---GGCCGGTTCGAG 364

RESULT 8

ABK93954
ID ABK93954 standard; DNA; 568 BP.

XX AC ABK93954;

XX XX 07-AUG-2003 (revised)

XX DT 27-AUG-2002 (first entry)

XX DE Corn zmICK1 EST DNA sequence.

XX KW Plant; ds; gene; inhibitor of cyclin dependent kinase; ICK.

XX OS Zea mays.

XX PN WO200228893-A2.

XX 11-APR-2002.
XX
XX
XX 29-JUN-2001; 2001WO-IB001492.
XX
XX
XX 14-JUL-2000; 2000US-0218471P.
XX 13-OCT-2000; 2000US-0241219P.
XX
XX (CROP-) CROPDESIGN NV.
XX
XX Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
PI Hatzfeld Y;
XX
XX WPI; 2002-471311/50.
DR P-PSDB; ABG65674.
XX
XX Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
PT to screen substrates, drugs or compounds which modulate ICK activity and
PT treat disorders characterized by an insufficient or excessive production
PT of ICK inhibitors.
XX
XX Claim 41; Page 125; 141pp; English.
XX
XX This invention relates to the DNA and protein sequences of novel isolated
CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
CC the invention may be used for treating disorders characterised by
CC insufficient or excessive production of an ICK inhibitor. The protein of
CC the invention may also be used to screen for naturally-occurring ICK
CC substrates, drugs or compounds which modulate ICK activity, as well as to
CC treat disorders characterised by insufficient or excessive production of
CC ICK protein, forms which have decreased or aberrant activity compared to
CC ICK wild type protein. The present sequence represents a cDNA molecule
CC encoding an inhibitor of cyclin dependent kinase (ICK) protein of the
CC invention. (Updated on 07-AUG-2003 to correct OS field.)
XX
XX Sequence 568 BP; 139 A; 149 C; 177 G; 102 T; 0 U; 1 Other;
SQ
Alignment Scores:
Pred. No.: 3.78e-15 Length: 568
Score: 399.50 Matches: 95
Percent Similarity: 81.90% Conservative: 0
Best Local Similarity: 81.90% Mismatches: 14
Query Match: 30.64% Indels: 7
DB: 6 Gaps: 7
US-09-993-808B-2 (1-256) x ABK93954 (1-568)
QY 146 GlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGlySerGlyProAspArg 165
Db 8 GGGGATCAGTC---CCGGATGTCGTCNCCGCGAGCAACTCGGGGAGCGTCCCGGACCGC 64
QY 166 GluArgArgGluThrThrPro---SerSerArgAlaHis---GlyGluLeuSerAspLeu 183
Db 65 GAGAGGAGAGAGACGACGACGCCATCGTCGAGCCGGCGCACGGGGCGAGCTCAGCGATCTG 124
QY 184 GluSerAspLeuAlaGly---HisLysThrGly---ProSerLeuProAla---AlaThr 200
Db 125 GAGTCGGATCTGTTGGGGCGGAGAAAGACTGGCTGCTCGTCCGCGGACAAACA 184
QY 201 ProAlaAlaGluLeuIleValProProAlaHisGluIleGlnGluPheAlaAlaAla 220
Db 185 TCGGCTCGGAGCTGATCGTGCCGCGAGCACAGGAGATCCAGGAATTCTTCGCGCGCC 244
QY 221 GluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyVal 240
Db 245 GAGGCGGCCCATGCCAAACGCTTTGCTTCCAAGTACAACTTCGACTTCGTCGCGCGCGTG 304
QY 241 ProLeuAspAlaGlyGlyArgPheGluTrpAlaProValValSerIle 256
Db 305 CCCCTCGACGCC---GGCCGGTTTCGAGTGGACGCCAGGGGTTCAGCATC 349
RESULT 9
ABK93981

ID ABK93981 standard; cDNA; 681 BP.
XX
XX AC ABK93981;
XX
XX 27-AUG-2002 (first entry)
XX
XX DE Full length Rice OsICK5 cDNA sequence.
XX
XX KW Plant; ds; gene; inhibitor of cyclin dependent kinase; ICK.
XX
XX OS Oryza sativa.
XX
XX PN WO200228893-A2.
XX PD 11-APR-2002.
XX
XX 29-JUN-2001; 2001WO-IB001492.
XX PF
XX 14-JUL-2000; 2000US-0218471P.
XX PR
XX 13-OCT-2000; 2000US-0241219P.
XX
XX PA (CROP-) CROPDESIGN NV.
XX
XX PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
PI Hatzfeld Y;
XX
XX DR WPI; 2002-471311/50.
XX DR P-PSDB; ABG65692.
XX
XX PT Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
PT to screen substrates, drugs or compounds which modulate ICK activity and
PT treat disorders characterized by an insufficient or excessive production
PT of ICK inhibitors.
XX
XX PS Claim 46; Page 139; 141pp; English.
XX
XX CC This invention relates to the DNA and protein sequences of novel isolated
CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
CC the invention may be used for treating disorders characterised by
CC insufficient or excessive production of an ICK inhibitor. The protein of
CC the invention may also be used to screen for naturally-occurring ICK
CC substrates, drugs or compounds which modulate ICK activity, as well as to
CC treat disorders characterised by insufficient or excessive production of
CC ICK protein, forms which have decreased or aberrant activity compared to
CC ICK wild type protein. The present sequence represents a cDNA molecule
CC encoding an inhibitor of cyclin dependent kinase (ICK) protein of the
CC invention
XX
XX SQ Sequence 681 BP; 102 A; 176 C; 307 G; 94 T; 0 U; 2 Other;
Alignment Scores:
Pred. No.: 4.55e-08 Length: 681
Score: 279.00 Matches: 98
Percent Similarity: 43.51% Conservative: 26
Best Local Similarity: 34.39% Mismatches: 71
Query Match: 21.40% Indels: 90
DB: 6 Gaps: 13
US-09-993-808B-2 (1-256) x ABK93981 (1-681)
QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
Db 1 ATGGGGAAG---AAGAAGACGCGACGGCGCGCGAGGAGGAGGCGCGGTGGTG 57
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyVal 40
Db 58 GTGGGC-----GGCGTCGTACGCGG----- 78
QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaValSer 60
Db 79 GCCCGCGTCACGGCGAGGAGGGTGGTGGCGAGCGCGAGAGGGTGTGGTTGGTGGGC 138
QY 61 AlaGlyGlyAspGlyGlySer-----CysTyrIleHisLeu 72

Db 139 CGTGGCGTGGCGGTGGCAGTGGCGGAGACGATGGCGAGGGCGGATCTATCTGCGTCTG 198
 QY 73 ArgSerArgMetLeuPheMetAlaProGlnProGlnProGlnProSerValAspSerValPro 92
 Db 199 CGGAGCAGGAGGCTG-----
 QY 93 ThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaGly 112
 Db 214 ---CCCTTCGTGGCGCGCGGTGGTGTCTGTCGCGGAGGAGGCGCTCGGTGATTCTG 270
 QY 113 LeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlnArgGly 132
 Db 271 GTGGCGGAGGCGGCTTCGTCTGTCGTCTGCGGCGGCGGTGGAATTGTTGGC----- 321
 QY 133 SerHisThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuValAsp 152
 Db 322 -----TGT-----TCTGTTGAGGAGGAGGCTATGGCCGAG 351
 QY 153 ValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThrPro 172
 Db 352 AAGNNGAGCGG-----ACGACGCGG 372
 QY 173 SerSerArg-----AlaHisGlyGluLeuSerAspLeuGluSerAsp----- 186
 Db 373 TCGAGCCCGCGCGCGCGGAGAGCGGACTCGAGCGAGTCAAACGAGGAGGCC 432
 QY 187 -----LeuAlaGlyHisLysThrGlyPro 194
 Db 433 AAGCAGCAAATGTGCGCGCGGAGTTCGACGACCTCAGCAGCTGCATTTACGCGGGA--- 489
 QY 195 SerLeuProAlaAlaThrProAlaAlaGluLeuIleValProPro-----AlaHisGlu 212
 Db 490 -----GCGACGACGAGGAGCTTCAGGATGATGGCACCGCGCGGCGGCGAGAG 540
 QY 213 IleGlnGluPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAlaSerLysTyr 232
 Db 541 ATCGAGGAGTTCCTCGCGCGTGGGAGAGGTCCGAGGCGGAGCGCTTCGCGCGCAAGTAC 600
 QY 233 AsnPheAspPheValArgGlyValProLeuAspAlaGly-----GlyArgPheGluTrp 250
 Db 601 AACTTCGACGTTGGTGGCGGCGGTGCGGCTCGACCGCGGCGGCGGCGGCGGTTTCGATGG 660
 QY 251 AlaProValValSer 255
 Db 661 ACCGCGTGGGCAGC 675

RESULT 10

ABK93982
 ID ABK93982 standard; DNA; 1073 BP.
 XX
 AC ABK93982;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE Rice OsICK5 genomic DNA clone.
 XX
 KW Plant; ds; gene; inhibitor of cyclin dependent kinase; ICK.
 XX
 OS Oryza sativa.
 XX
 PN WO200228893-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 29-JUN-2001; 2001WO-IB001492.
 XX
 PR 14-JUL-2000; 2000US-0218471P.
 PR 13-OCT-2000; 2000US-0241219P.
 XX
 PA (CROP-) CROPDESIGN NV.
 XX
 PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;

PI Hatzfeld Y;
 XX
 DR WPI; 2002-471311/50.
 XX
 PT Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
 PT to screen substrates, drugs or compounds which modulate ICK activity and
 PT treat disorders characterized by an insufficient or excessive production
 PT of ICK inhibitors.
 XX
 PS Claim 46; Page 140; 141pp; English.
 XX
 CC This invention relates to the DNA and protein sequences of novel isolated
 CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
 CC the invention may be used for treating disorders characterised by
 CC insufficient or excessive production of an ICK inhibitor. The protein of
 CC the invention may also be used to screen for naturally-occurring ICK
 CC substrates, drugs or compounds which modulate ICK activity, as well as to
 CC treat disorders characterised by insufficient or excessive production of
 CC ICK protein, forms which have decreased or aberrant activity compared to
 CC ICK wild type protein. The present sequence represents a cDNA molecule
 CC encoding an inhibitor of cyclin dependent kinase (ICK) protein of the
 CC invention
 XX

SQ Sequence 1073 BP; 172 A; 287 C; 417 G; 197 T; 0 U; 0 Other;

Alignment Scores:
 Pred. NO.: 1.01e-05 Length: 1073
 Score: 242.00 Matches: 117
 Percent Similarity: 38.20% Conservative: 27
 Best Local Similarity: 31.03% Mismatches: 90
 Query Match: 18.56% Indels: 144
 DB: 6 Gaps: 19

US-09-993-808B-2 (1-256) x ABK93982 (1-1073)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaValGlu 20
 Db 1 ATGGGGAAG---AAGAAGAAGCGCGACGCGCGCGCGCGCGGAGGAGGCGGGTGGTG 57
 QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal 40
 Db 58 GTCGCG-----GGCGTCCGTACGCGG----- 78
 QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaValSer 60
 Db 79 GCCCGGTACCGCGGAGGAGGTTGGTGGCGAGCGCGGAGGAGGAGGAGGAGGAGGAG 138
 QY 61 AlaGlyGlyAspGlyGlySer-----CystyrIleHisLeu 72
 Db 139 CGTGGCGGTGGCGGTGGCGGAGTGGCGGAGACGATGGCGGAGGCGGATCTATCTGCGTCTG 198
 QY 73 ArgSerArgMetLeu---PheMetAlaProPro-----GlnProGlnPro 86
 Db 199 CGGAGCAGGAGGCTGCCCTTCGTGGCGGCGCGCGGTGGTGTCTGCGCGGAGGAGGAGGAG 258
 QY 87 SerValAspSerValProThrProValGluAlaAlaAsp----- 99
 Db 259 CTCGGTATTCTGGTG-----GCGGAGGCGGCTTCGTCTGCGTCTGCGCGGCGGCGGTG 309
 QY 100 -----GlyAlaAlaGlyGlnGlyAla----- 107
 Db 310 GAATTGTTGGGCTGTTCTGGTGGAGGAGGAGGCTATGGCCGAGAGGAGGAGGAGGAGGAG 369
 QY 108 -----AlaLeuAlaAlaGlyLeuSerArg---CysSerSerThrAlaSerSerVal 123
 Db 370 TAGAATTCCTCCCGGCTCGAGTCTCGATCGCCCGCTTCATCTCTTGTCTGAATGATGC 429
 QY 124 AsnLeuGlyLeuGly-----GlyGlnArgGlySerHisThrCys 136
 Db 430 GGCTTGGGATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 488
 QY 137 Arg-----SerTyrAspAlaAlaGluAlaGlyGlyAspHisVal----- 149

Db 489 AGAGCTCCGTCGGCGACTCCGGCTCGCGCGCGAGAGGTGATCGAGCTCCTCTCCACGCG 548
 QY 150 -----LeuValAspValSerAlaAlaSerAsnSerGlySerGlyPro 163
 Db 549 TTCTTGTCTTGTCTTGACATGATTAATTACAACCGCGGTCTCTCAATTGAATTATCGCA 608
 QY 164 AspArgGluArgArgGluThrThrProSerSerArg-----AlaHisGlyGluLeu 180
 Db 609 ATTCAATCCAGGAGCGCGACGACGCGCTCGAGCGCGCGCGCGCGGAGCGGACTCG 668
 QY 181 SerAspLeuGluSerAsp-----186
 Db 669 AGCGACGCGGAGTCAAAACGAGGAGCGCAAGCAATGTGCGCGCGGAGTTCGACGACC 728
 QY 187 ---LeuAlaGlyHisLysThrGlyThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeu 205
 Db 729 TCAGCAGCTGCATTTCACGGGGA-----GCGACGACGAGGAGCTTCAGGATG 776
 QY 206 IleValProPro-----AlaHisGluIleGlnGluPhePheAlaAlaGluAlaAla 223
 Db 777 ATGGCACCG 836
 QY 224 GlnAlaLysArgPheAlaSerLys-----231
 Db 837 GAGGCGGAGCGCTTCGCCGCCAAGTGAAGTGTGTCATCACATATTGTCTCGTCCGTGCGT 896
 QY 231 -----231
 Db 897 GTCGTACATATCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 956
 QY 232 -----TyrAsnPheAspPheValArgGlyVal 240
 Db 957 GGAGCTGATTAACTGCGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1016
 QY 241 ProLeuAspAlaGly-----GlyArgPheGluTrpAlaProValValSer 255
 Db 1017 CCGCTCGACG 1067

RESULT 11

ABK93971
 ID ABK93971 standard; cDNA; 1067 BP.

XX AC ABK93971;

XX 27-AUG-2002 (first entry)

XX DE Full length Rice OsICK4 cDNA sequence.

XX KW Plant; ss; primer; PCR; inhibitor of cyclin dependent kinase; ICK.

XX OS Oryza sativa.

XX PN WO200228893-A2.

XX PD 11-APR-2002.

XX PF 29-JUN-2001; 2001WO-IB001492.

XX PR 14-JUL-2000; 2000US-0218471P.

XX PR 13-OCT-2000; 2000US-0241219P.

XX PA (CROP-) CROPDESIGN NV.

XX PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;

XX PI Hatzfeld Y;

XX DR WPI; 2002-471311/50.

XX DR P-PSDB; ABG65691.

XX PT Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used

PT to screen substrates, drugs or compounds which modulate ICK activity and

PT treat disorders characterized by an insufficient or excessive production

PT of ICK inhibitors.

XX PS Claim 46; Fig 2; 141pp; English.
 XX CC This invention relates to the DNA and protein sequences of novel isolated
 CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
 CC the invention may be used for treating disorders characterised by
 CC insufficient or excessive production of an ICK inhibitor. The protein of
 CC the invention may also be used to screen for naturally-occurring ICK
 CC substrates, drugs or compounds which modulate ICK activity, as well as to
 CC treat disorders characterised by insufficient or aberrant activity production of
 CC ICK protein, forms which have decreased or aberrant activity compared to
 CC ICK wild type protein. The present sequence represents a PCR primer
 CC specific for an inhibitor of cyclin dependent kinase (ICK) gene used in
 CC the examples of the specification
 XX SQ Sequence 1067 BP; 247 A; 292 C; 300 G; 226 T; 0 U; 2 Other;

Alignment Scores:
 Pred. No.: 2.55e-05 Length: 1067
 Score: 235.00 Matches: 82
 Percent Similarity: 41.15% Conservative: 25
 Best Local Similarity: 31.54% Mismatches: 73
 Query Match: 18.02% Indels: 80
 DB: 6 Gaps: 12

US-09-993-808B-2 (1-256) x ABK93971 (1-1067)
 QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
 Db 39 ATGGGCAAGTACATGCGCAAGGCCAAGGTGGTGTCTCCGGCGAGGTGGTGGCGCGCC 98
 QY 21 ValThrGlnVal-----ValGlyValArgThrArgSerArgSerAlaAlaAla 36
 Db 99 GTCATGGAGCTCG 158
 QY 37 ThrGlyGlyValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAla 56
 Db 159 -----CAGAAGAGG-----167
 QY 57 AlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMet 76
 Db 168 -----CAGGGCGGGGAG-----TACCTCGAGTCCAGGAGCGCGCAGG 203
 QY 77 LeuPheMetAlaProProGlnProGlnProSerValAspSerValProThrProValGlu 96
 Db 204 CTCGAGAAGCTCCCT 245
 QY 97 AlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCys 116
 Db 246 AGGGCGACGGCTGGGCT-----GGACTGCT 272
 QY 117 SerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlnArgGlySerHisThrCys 136
 Db 273 GATCGACGGCGACG-----287
 QY 137 ArgSerTyrAspAlaAlaGluAla-----GlyGlyAspHisValLeuValAsp 152
 Db 288 -----GAGAGCGCGGAGCGGAGGTGTCTCGTTCGGGGGAGAACGTCCTC---GAG 335
 QY 153 ValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThrPro 172
 Db 336 CTGGAGGCCATG-----GAAAGGAATACCAGGAGGAGACGACCT 374
 QY 173 SerSer-----ArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHis 190
 Db 375 TGCAGCTTGATCAGGGACCCCGATACGATTAGCACCCCTGGATCTACCACAGGCGCAGC 434
 QY 191 LysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeuValProProAla 210
 Db 435 CACTCGAGTTCTCATTCGAAGGTGCAACACCCCGTGGCCCAACATTATTCCAGCATCA 494
 QY 211 HisGluIleGlnGluPheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAlaSer 230

Db 495 GCAGAGCTGGAAGCGTTCTTCGCGCCGAGAGCAACGGCAACGACAGGCTTTTCATCGAC 554
QY 231 LysTyrAsnPheAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGluTrp 250
Db 555 AAGTATAACTTTGATCCTGTGAATGACTGCCCTCTTCCC-----GGCCGRTTTTGAATGG 608

RESULT 12

AAA95292
ID AAA95292 standard; cDNA; 1169 BP.

XX AC AAA95292;

XX 17-JAN-2001 (first entry)

XX Wheat cyclin-dependent kinase inhibitor coding sequence #2.

XX Wheat; cyclin-dependent kinase inhibitor; cell cycle; cell division;
KW CDKI; cell growth; herbicide; ss.

XX Triticum aestivum.

XX Key Location/Qualifiers
FT CDS 141..713
FT /*tag= a
FT /product= "CDKI"

XX WO200060087-A2.

XX 12-OCT-2000.

XX 06-APR-2000; 2000WO-US009106.

XX 07-APR-1999; 99US-0128192P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Klein TM, Weng Z, Cahoon RE;

XX WPI; 2000-679375/66.

XX P-PSDB; AAB27258.

XX Cyclin dependent kinase inhibitor sequences, useful for identifying
PT herbicides and plant growth inhibitors.

XX Claim 2; Page 52-53; 58pp; English.

XX The present sequence is the coding sequence for the wheat cyclin-
CC dependent kinase inhibitor (CDKI). It was isolated by searching a wheat
CC root cDNA library for sequences similar to those encoding the CDKI from
CC Chenopodium rubrum, Caenorhabditis elegans and Arabidopsis thaliana. CDKI
CC is involved in the cell cycle, and may promote or inhibit cell division
CC and growth. The coding sequence and the protein it encodes are useful in
CC the production of transgenic plants which produce increased or decreased
CC amounts of the CDKI protein, in the identification of herbicides, in
CC genetic and physical mapping and in the isolation of the CDKI gene in
CC other organisms

XX Sequence 1169 BP; 263 A; 338 C; 333 G; 235 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0.000238 Length: 1169
Score: 219.00 Matches: 87
Percent Similarity: 41.09% Conservative: 19
Best Local Similarity: 33.72% Mismatches: 72
Query Match: 16.79% Indels: 80
DB: 3 Gaps: 15

US-09-993-808B-2 (1-256) x AAA95292 (1-1169)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaGlyAlaGluValAlaValGlu 20
Db 141 ATGGGCAAGTACATGCGCAAGCCCAAG-----GTCTCCGGCGAGGTGGCGTCATGGAG 194

QY 21 ValThrGlnVal---ValGlyValArgThrArgSerArgSerAlaAlaThrGlyGly 39
Db 195 GTCCGCCGCCGCCGCTAGGGTCCGACCCCGCGCACGAGCGCTCGCG----- 242
QY 40 ValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaVal 59
Db 243 -----ATGCAGAGGCGCGCGAGGG-----GCGGCGGTG 272
QY 60 SerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMet 79
Db 273 GCCAAG-----GACCAGGGGAG---TACCTGGAGCTCAGAGTCGGAAGTCGAGAAG 323
QY 80 AlaProProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAsp 99
Db 324 CTGCCCCCGCGCGCGCG----- 341
QY 100 GlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThr 119
Db 342 ---GCGGCGAGGAGGCGCGCGCG----- 365
QY 120 AlaSerSerValAsnLeuGlyLeuGlyGlyGlnArgGlySerHisThrCysArgSerTyr 139
Db 366 GCGAGCGGTGTCGAG----- 380
QY 140 AspAlaAlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSer 159
Db 381 -----GCCGAGGCGCGAGGCGCGAGGTGCTCTCGGTGAGAACGTCGAGTCGGAG 434
QY 160 GlySerGlyProAspArgGluArgGluThrThrProSerSer-----ArgAlaHis 177
Db 435 GCCATGGGG-----AGGGGTACAGGAGGAGACGCGCCCTGCAGCTTGATTAGGACTCG 488
QY 178 GlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSer----- 195
Db 489 GGAACGATAAGC-----ACTCCTGGATCCACAAAGACCGAGCCACTCG 533
QY 196 -----LeuProAlaAlaThrProAlaAlaGluLeuValProProAlaHisGlu 212
Db 534 AATTCCTCATCGCAGGGTGCAAGCTCCAGCGCGCCAT---ATTATCCATGTTTCAGCAGAG 590
QY 213 IleGlnGluPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAlaSerLysTyr 232
Db 591 ATGAATGAGTTCTTCTCTGTCTGGGAGCAACCGCAACAGCAAGCCTTCATTGACAAGTAC 650
QY 233 AsnPheAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGluTrp 250
Db 651 AACTTTGATCCTGTGAACGACTGTCTCTCCCA-----GGCCGATACGAGTGG 698

RESULT 13

AAD40771

ID AAD40771 standard; cDNA; 797 BP.

XX AAD40771;

XX 30-OCT-2002 (first entry)

XX Arabidopsis thaliana ICN7 full-length cDNA.

XX Plant development; cyclin-dependent kinase; CDK inhibitor; ICK1; ICK2;
KW ICN2; ICN6; ICN7; ICN8; ICNK; morphogenesis; maturation; enlargement;
KW plant breeding; growth; gene; ss.

XX Arabidopsis thaliana.

XX Location/Qualifiers
FH 10..678
FT /*tag= a
FT /product= "ICN7 full-length protein"

XX WO200250292-A2.

XX 27-JUN-2002.

XX
AC
ABV74604;
XX

Qy 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAAlaGlyAlaGluValAlaAlaValGlu 20

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2004, 09:59:54 ; Search time 85 Seconds
(without alignments)
1671.382 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 1304
Sequence: 1 MGKYMKCRGAAGAEVAE.....VRGVPLDAGGFEPWAPVWSI 256

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues
Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US0993808/runat_01102004_171703_13045/app_query.fasta_1.455
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=Dits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0993808@cgn_1_1_105@runat_01102004_171703_13045 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PTCUS COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length DB | ID | Description |
|------------|-------|---------------|-----------|----|---------------------------------------|
| C 1 | 152.5 | 11.7 | 4403765 | 3 | US-09-103-840A-2 Sequence 2, Appli |
| 2 | 141.5 | 10.9 | 4403765 | 3 | US-09-103-840A-2 Sequence 2, Appli |
| 3 | 141.5 | 10.9 | 4411529 | 3 | US-09-103-840A-1 Sequence 1, Appli |
| C 4 | 140.5 | 10.8 | 4411529 | 3 | US-09-103-840A-1 Sequence 1, Appli |
| C 5 | 134 | 10.3 | 2561 | 4 | US-09-616-289-48 Sequence 48, Appl |
| C 6 | 130.5 | 10.0 | 1845 | 4 | US-09-252-991A-3639 Sequence 3639, Ap |
| C 7 | 130.5 | 10.0 | 12425 | 4 | US-09-616-289-50 Sequence 50, Appl |
| C 8 | 129.5 | 9.9 | 4848 | 4 | US-08-955-957A-1 Sequence 1, Appli |
| C 9 | 129.5 | 9.9 | 4848 | 4 | US-08-955-957A-4 Sequence 4, Appli |
| C 10 | 129.5 | 9.9 | 4848 | 4 | US-08-955-957A-6 Sequence 6, Appli |
| 11 | 126.5 | 9.7 | 33529 | 3 | US-09-144-085-3 Sequence 3, Appli |
| 12 | 125.5 | 9.6 | 1143 | 4 | US-09-457-568-7 Sequence 7, Appli |

| | | | | | |
|------|-------|-----|-------|---|--|
| 13 | 125.5 | 9.6 | 1143 | 4 | US-09-457-646-7 Sequence 7, Appli |
| 14 | 125.5 | 9.6 | 1143 | 4 | US-08-902-572-5 Sequence 5, Appli |
| 15 | 125.5 | 9.6 | 1143 | 4 | US-09-516-065-7 Sequence 7, Appli |
| C 16 | 125 | 9.6 | 861 | 4 | US-09-347-613C-8 Sequence 8, Appli |
| 17 | 125 | 9.6 | 2481 | 4 | US-09-894-998A-35 Sequence 35, Appl |
| 18 | 125 | 9.6 | 13987 | 2 | US-08-804-227C-13 Sequence 13, Appl |
| 19 | 123.5 | 9.5 | 807 | 4 | US-09-403-768-1 Sequence 1, Appli |
| 20 | 123.5 | 9.5 | 1983 | 4 | US-09-252-991A-16300 Sequence 16300, A |
| 21 | 123 | 9.4 | 3915 | 4 | US-09-489-039A-6278 Sequence 6278, Ap |
| C 22 | 123 | 9.4 | 53526 | 3 | US-08-658-136-2 Sequence 2, Appli |
| C 23 | 123 | 9.4 | 53577 | 3 | US-08-658-136-1 Sequence 1, Appli |
| 24 | 122.5 | 9.4 | 34094 | 4 | US-09-292-034-1 Sequence 1, Appli |
| C 25 | 122.5 | 9.4 | 43280 | 2 | US-08-804-227C-1 Sequence 1, Appli |
| C 26 | 121.5 | 9.3 | 1070 | 4 | US-09-470-443-7 Sequence 7, Appli |
| 27 | 121.5 | 9.3 | 1659 | 3 | US-09-083-351-3 Sequence 3, Appli |
| 28 | 121.5 | 9.3 | 1659 | 3 | US-09-083-352-3 Sequence 3, Appli |
| 29 | 121.5 | 9.3 | 3946 | 3 | US-09-083-351-1 Sequence 1, Appli |
| 30 | 121.5 | 9.3 | 3946 | 3 | US-09-083-352-1 Sequence 1, Appli |
| 31 | 121 | 9.3 | 2472 | 4 | US-09-252-991A-8197 Sequence 8197, Ap |
| 32 | 121 | 9.3 | 43280 | 2 | US-08-804-227C-1 Sequence 1, Appli |
| 33 | 120 | 9.2 | 1098 | 4 | US-09-457-568-9 Sequence 9, Appli |
| 34 | 120 | 9.2 | 1098 | 4 | US-09-457-646-9 Sequence 9, Appli |
| 35 | 120 | 9.2 | 1098 | 4 | US-08-902-572-7 Sequence 7, Appli |
| 36 | 120 | 9.2 | 1098 | 4 | US-09-516-065-9 Sequence 9, Appli |
| C 37 | 120 | 9.2 | 1323 | 4 | US-09-252-991A-3022 Sequence 3022, Ap |
| C 38 | 120 | 9.2 | 1965 | 4 | US-09-252-991A-9230 Sequence 9230, Ap |
| C 39 | 120 | 9.2 | 2805 | 4 | US-09-252-991A-9064 Sequence 9064, Ap |
| 40 | 120 | 9.2 | 4190 | 3 | US-08-938-291A-2 Sequence 2, Appli |
| 41 | 120 | 9.2 | 4190 | 4 | US-09-589-619-2 Sequence 2, Appli |
| 42 | 119.5 | 9.2 | 507 | 4 | US-09-105-470B-28 Sequence 28, Appl |
| C 43 | 119 | 9.1 | 702 | 4 | US-09-252-991A-8108 Sequence 8108, Ap |
| C 44 | 119 | 9.1 | 1533 | 4 | US-09-252-991A-12896 Sequence 12896, A |
| C 45 | 119 | 9.1 | 1854 | 4 | US-09-894-844-103 Sequence 103, App |

ALIGNMENTS

RESULT 1
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 35.4 Length: 4403765
Score: 152.50 Matches: 75
Percent Similarity: 35.91% Conservative: 18
Best Local Similarity: 28.96% Mismatches: 100
Query Match: 11.69% Indels: 66
DB: 3 Gaps: 10

US-09-993-808B-2 (1-256) x US-09-103-840A-2 (1-4403765)

; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 294 Length: 4411529
Score: 141.50 Matches: 72
Percent Similarity: 38.96% Conservative: 18
Best Local Similarity: 31.17% Mismatches: 100
Query Match: 10.85% Indels: 42
DB: 3 Gaps: 9

US-09-993-808B-2 (1-256) x US-09-103-840A-1 (1-4411529)

QY 6 ArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGluValThrGln----- 23
Db 928765 CGCAGTGCGCGGTAGCGCTGCCGCC--GTTGCCCGCGTTGGCGCGGTTAAACCCCGAC 928821
QY 24 -----ValValGlyValArgThr---ArgSerArgSerAlaAlaThrGly 38
Db 928822 TCCGGTGCGCGGTTGCCCGGTTGCCACCTGCCCGACGAATCCGAAGCGGTACCCGCC 928881
QY 39 GlyValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaAla 58
Db 928882 GGCACCGCC--GCTGCCCGCGGTACCAACCGAAGCCCGCGCGGTGCCACCGCGGCC 928938
QY 59 ValSerAlaGlyGlyAspGlyGly-SerCysTyrIleHisLeuArgSerArgMetLeuPh 78
Db 928939 GCCCCACGCGCCAGCAGCGCGGTCCCG-----CTGCC 928971
QY 78 eMetAlaProGlnProGlnProSerValAspSerValProThrProValGluAlaAl 98
Db 928972 TCCGGCGCGCGGTTGCCCGCGGTTCGGTGGCGGCTCCGCGCATGGCCGCGCTAC 929090
QY 98 aAspGlyAlaAlaGlyGlnGln-----GlyAlaAlaLeuAlaAlaGlyLe 113
Db 929031 CGATGCCGCGCGCATCAATCCAGAGGCGATCGCCGCGGTCCGCGCATGGCCGCGTAC 929090
QY 113 uSerArgCysSerSerThrAlaSerSerValAsnLeuGlyGlyGlnArgGlyse 133
Db 929091 CAGCCGAAGCGGCGCGCGCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 929150
QY 133 rHisThrCysArgSerTyrAspAlaAlaGluAlaGlyGly----- 146
Db 929151 CCAATCCG 929210
QY 147 -----AspHisValLeuValAspValSerAlaAlaSerAsnSerGlySerGlyPro-- 163
Db 929211 CGCCG 929270
QY 164 -----AspArgGluArgGluThrThrProSerSerArgAlaHisGlyGluLe 180
Db 929271 CGGTGCCG 929328
QY 180 uSerAspLeuGluSerAsp-LeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaT 200
Db 929329 -----GCCGAACCGGATGAGTCGGGGGCAACAGCGAAGCGCGCGCGCGCGCGCG 929381
QY 200 hrProAlaAlaGluLeuIleValProPro 209
Db 929382 CGCCGTTCCAGTGGTAGCTCGCGCACCA 929410

RESULT 4
US-09-103-840A-1/c

; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:

Pred. No.: 356 Length: 4411529
Score: 140.50 Matches: 81
Percent Similarity: 38.46% Conservative: 19
Best Local Similarity: 31.15% Mismatches: 104
Query Match: 10.77% Indels: 57
DB: 3 Gaps: 11

US-09-993-808B-2 (1-256) x US-09-103-840A-1 (1-4411529)

QY 6 ArgLysCysArg-GlyAlaAlaGlyAlaGluValAlaAlaValGluValThrGlnValVa 25
Db 3948721 CGCCTTGCCGCGGTGCCCGCGGTGCCCGCGGTGCCCGCGGTGGTGGCGTCTGCTAGTCCGT 3948662
QY 25 lGlyVal-----ArgThrArgSerArgSerAlaAlaThrGlyGlyValAlaLysVa 43
Db 3948661 CGGTGTTAGCCGCGCTTTGCCGCGGTGCCCGCGGTGCCCGCGGTGCCCGCGGTGAAGT 3948602
QY 43 lAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaValSerAlaGlyGl 63
Db 3948601 TGGCCCC-----GCCGCTACCGCGCGC---CCCGCGCTGCCCGCGGTGCCCGCGG 3948554
QY 63 yAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAla-ProProG 83
Db 3948553 TGCCGCGGTGCCGCGCTTTGCCGCGGTGCCCGCGCGCGGTGCCGTGCCCGCGGTGCC 3948494
QY 83 lnProGlnProSerValAspSerValPro-----ThrProValGluAlaAla 98
Db 3948493 CCCCCTGCCCG 3948435
QY 98 laAsp-----GlyAlaAlaGlyGlnGlnGlyAlaAla 108
Db 3948434 CTGATGCCAGCACCCCG 3948375
QY 108 laLeuAlaAlaGlyLeuSerArgCysSerSerThrAlaSerSer----- 122
Db 3948374 TTGTTGCCCG 3948315
QY 123 -----ValAsnLeuGlyLeuGlyGlnArgGlys 133
Db 3948314 AACCCGCGTCCG 3948255
QY 133 erHisThrCysArgSerTyrAspAlaAlaGlu-----AlaGlyGlyAspHisV 149
Db 3948254 CCGTACCGCGCTTGGCG 3948196
QY 149 alLeuValAspValSerAlaAlaSerAsnSerGlySerGlyProArgGluArgArg 169
Db 3948195 -----GCCGCGGTGCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3948168
QY 169 luThrThrPro-----SerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAspL 187


```
QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaAlaValSer 60
Db 1130 GCGACAGCATCGCTCGTCCAGGCGTGTGCGCAGGTGCGCGCAGCGGT----- 1077

QY 61 AlaGlyGlyAspGly-GlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAl 80
Db 1076 GGTGGCGATCAGCGCGCGTGGCTCGCCATCGCGCGAGTCTGGATACCGTAGC 1017

QY 80 aProProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAspGl 100
Db 1016 C-----GCCCTGGCAGCGACCCAGAGAGCCTTCGCGATCGGCGG 975

QY 100 yAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerThrAl 120
Db 974 CGTAGCGGCCACACAGGTGCGCGTGGCGAGAACGAGCGGAGCCCTGCCAGGTGTGT 915

QY 120 aSerSerValAsnLeuGlyLeuGlyGlnArgGlySerHisThrCysArgSerTyrAs 140
Db 914 CCGGCGCGCGAAT---GGTGAGGGTGGT-----GGCTTCTTCGA 879

QY 140 pAlaAlaGluAlaGlyGlyAsp-----HisValLeu----- 150
Db 878 TCAGGTACATGCGGTGGCGATGTCCAGTGTCTCGGGTGCACGTCTGGGTTCGACCG 819

QY 151 -ValAspValSerAlaAlaSerAsnSerGlySerGly----- 162
Db 818 GATCGCGGTGGCGCGGAGCCGAGCAGCATGCCGCGTCCGGTTCAGGTAGAGGATT 759

QY 163 -----ProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyGlu-- 179
Db 758 CGTCGAGGCTGACACAGCATCGGCCAGTCTGTGGCAGTGCATGCCGCGGCGGGAAGA 699

QY 180 -----LeuSerAspLeuGluSerAspLeuAlaGlyHisLysTh 192
Db 698 TGAAGGCTGAGCGCGCTTGGGCTGCAGGCCAGCGGGCGGAC---GCCGCGCAGCCCGG 642

QY 192 rGlyProSerLeuProAlaAlaThr----- 200
Db 641 CGATGGCGTGCACACAGGCGCTGCGCGTGTGACACAGCACCGCGCGGGTAGTTGCCGG 582

QY 201 -----ProAlaAlaGluLeuIleValProProAlaHisGluIleGlnGl 215
Db 581 CGTCGACGGGACTTCCAGGCGCC-----GTGACCCGCGG---GATCTCCAGCG 534

QY 215 uPhePheAlaAlaAlaGlu-----AlaAlaGlnAlaLysArgPheAlaSerLy 231
Db 533 CTTCGTGGTTGCACAGGACCTGCCCTGTTGCGCGGATGCGCGCAGGTAGCC----- 479

QY 231 sTyrAsnPheAspPheValArgGlyValProLeuAspAlaGly 245
Db 478 -----CTGGTGACGGCGCTCGGTATCGATGTCGGC 449
```

RESULT 7

```
US-09-616-289-50/c
; Sequence 50, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
```

```
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 50
; LENGTH: 12425
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-616-289-50

Alignment Scores:
Pred. No.: 0.923 Length: 12425
Score: 130.50 Matches: 69
Percent Similarity: 30.39% Conservative: 17
Best Local Similarity: 24.38% Mismatches: 71
Query Match: 10.01% Indels: 126
DB: 4 Gaps: 9
```

US-09-993-808B-2 (1-256) x US-09-616-289-50 (1-12425)

```
QY 8 CysArgGlyAlaAlaGlyAlaGluValAlaAlaValGluValThrGlnValValGlyVal 27
Db 3496 TGTGCGGCGCGCGCGCTGTGGC-----GGCGGCGGC 3464

QY 28 ArgThrArgSerArgSerAlaAlaAlaThrGlyGlyValAlaLysValAlaProArgArg 47
Db 3463 AGCGCGGCTCCCGCGCGCGACGCGGGCGGGGCGG----- 3425

QY 48 LysArgAlaProAlaGlyGluProAlaAlaAlaValSerAlaGlyGlyAspGlyGlySer 67
Db 3424 CGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3368

QY 68 CysTyrIleHisLeuArgSerArgMetLeuPheMetAlaProProGlnProGlnProSer 87
Db 3368 ----- 3368

QY 88 ValAspSerValProThrProValGluAlaAlaAspGlyAlaAla----- 102
Db 3367 -----GCGCGCAGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3335

QY 103 -----GlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCys 116
Db 3334 CCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3275

QY 117 SerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlnArgGlySerHisThrCys 136
Db 3274 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3215

QY 137 ArgSerTyrAspAlaAlaGluAlaGlyGlyAsp----- 147
Db 3214 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3155

QY 147 ----- 147
Db 3154 GGGGTGGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3095

QY 148 -----HisValLeuValAspValSerAlaAlaSer 157
Db 3094 TTGTAGCTGACCCGAGCACCGCGCGCGCTGCTGGATCAGTTCTCGAGCTCGGCGCGCGGTG 3035

QY 158 AsnSerGlySerGlyProAspArgGluArg----- 167
Db 3034 CGCTCCGCGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2975

QY 168 -----ArgGluThrThrProSer-----SerArgAlaHis-----Gly 178
Db 2974 CGCGCTTTCGCGCGAGCGCAGCGAGTCTGATGTTGTCAGGATCCACTCTTGGTAGTGGCGG 2915

QY 179 GluLeuSerAspLeuGluSerAspLeuAla----- 188
Db 2914 GAAGCGCGGACGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2855

QY 189 GlyHisLysThrGlyProSerLeuProAlaAlaThr-ProAlaAlaGluLeuIleValPr 208
```

Db 2854 GGTAGGCGGGGCCCCCGCATGCCTCCCGCCCGCGGCTGCACCGTGGCGGTGCC 2795
QY 208 oProAla 210
Db 2794 TCCCTCC 2788

RESULT 8

US-08-955-957A-1/c
; Sequence 1, Application US/08955957A
; Patent No. 6312920
; GENERAL INFORMATION:
; APPLICANT: DeHoff, Bradley S.
; APPLICANT: Rostock Jr., Paul R.
; TITLE OF INVENTION: SAM Operon
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/955,957A
; FILING DATE:

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: P-10162
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4848 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 986..2209
FEATURE:
NAME/KEY: CDS
LOCATION: 2241..3341
US-08-955-957A-1

Alignment Scores:
Pred. No.: 0.316 Length: 4848
Score: 129.50 Matches: 65
Percent Similarity: 36.11% Conservative: 26
Best Local Similarity: 25.79% Mismatches: 91
Query Match: 9.93% Indels: 70
DB: 4 Gaps: 12

US-09-993-808B-2 (1-256) x US-08-955-957A-1 (1-4848)

QY 26 GlyValArgThrArgSerArgSerAlaAlaAlaThrGlyValAlaAlaLysValAlaPro 45
Db 3404 GGGGAAGAACTCGTGGAGAGAGAGCGCGCTCCCGAGGATCTCGCGCAGGTGG 3345
QY 46 ArgArgLysArgAlaProAlaGlyGluProAlaAlaAlaValSerAlaGlyGlyAspGly 65
Db 3344 TCCTACCGGACGCTCCCGGTCGGCCCGCGCGCGCGCGCGGCGGCGGATGACGGCGGC 3285

QY 66 GlySerCysTyrIle-----His 71
Db 3284 GGCCCGTCCCGGGCGTCCGGCGGTGGCTCGCGCCGGGAGCCGGGACAGGGCCGGTGCAT 3225
QY 72 LeuArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAspSerVal 91
Db 3224 GTGCGG-----CCGCCAGCCCGAGGGCGAGCCGCCGCCGCGG 3186
QY 92 ProThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAla 111
Db 3185 TCCGCGCCGTAC-----GGG 3171
QY 112 GlyLeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlnArg 131
Db 3170 TCCCTCGCGGTGGCGAGCGGATCCCGCGGATCGGTCCGTAGGTCTGCGGGCGCGACCGT 3111
QY 132 GlySerHisThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuVal 151
Db 3110 CCCCAGCGCCCG-----CGCCCGCAGGACGCGCCAGCCCGCGCGCGGAGAC----- 3063
QY 152 AspValSerAlaAlaSerAsnSerGly-----SerGlyProAspArgGluArgGlu 169
Db 3062 GATGAAAGCG 3012
QY 170 ThrThrProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAlaGly 189
Db 3011 ACCGATCCCGGTGGATCGGATCGACACCGGGGACATCGGGGAC-----CGCGG 2967
QY 190 HisLys-ThrGly-----ProSerLeuProAlaAlaThrProAlaAlaGluLeuLeuVa 207
Db 2966 CACGTCAGCGGTGGTCCCG 2907
QY 207 1-----ProProAlaHisGluIleGlnGluPhePheAlaAlaAlaGluAlaAlaGl 224
Db 2906 GACCAGGCG 2847
QY 224 nAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyValProLeuAspAl 244
Db 2846 CAGC-----GCGGCCTCGTACTCGTTG----- 2825
QY 244 aGlyGlyArgPheGluTrpAlaProValValSer 255
Db 2824 ----GTGAAGACCCAGTGGCGCGCGCGTGCACCAGC 2795

RESULT 9

US-08-955-957A-4/c
; Sequence 4, Application US/08955957A
; Patent No. 6312920
; GENERAL INFORMATION:
; APPLICANT: DeHoff, Bradley S.
; APPLICANT: Rostock Jr., Paul R.
; TITLE OF INVENTION: SAM Operon
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/955,957A
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: P-10162

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-3334
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 4848 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 3338..4255
US-08-955-957A-4

Alignment Scores:
Pred. No.: 0.316 Length: 4848
Score: 129.50 Matches: 65
Percent Similarity: 36.11% Conservative: 26
Best Local Similarity: 25.79% Mismatches: 91
Query Match: 9.93% Indels: 70
DB: 4 Gaps: 12

US-09-993-808B-2 (1-256) x US-08-955-957A-4 (1-4848)

QY 26 GlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyValAlaLysValAlaPro 45
Db 3404 GGGGAAGAACTCGTGGAGAGAGAGCGCGCTGCCGAGGATCTCGCGAGGTTGG 3345

QY 46 ArgArgLysArgAlaProAlaGlyGluProAlaAlaAlaValSerAlaGlyGlyAspGly 65
Db 3344 TCCTACCGGCAGCCTCCCGCGGTCCGCCCGCGCGCGCGCGCGCGGCGGATGATGACGGCGGC 3285

QY 66 GlySerCysTyrIle-----His 71
Db 3284 GGCCCGTCCGGCGGTCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCAT 3225

QY 72 LeuArgSerArgMetLeuPheMetAlaProGlnProGlnProSerValAspSerVal 91
Db 3224 GTGCGG-----CGGCCAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3186

QY 92 ProThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAla 111
Db 3185 TCCGCGCGGTAC-----GCG 3171

QY 112 GlyLeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlnArg 131
Db 3170 TCCCTCGCGGTGGGAGCGGATCCGCCGATCGTCCGTTAGTCTCGGGGCGGACCGT 3111

QY 132 GlySerHisThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuVal 151
Db 3110 CCCCAGGCGCG--CGCCGCGAGGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 3063

QY 152 AspValSerAlaAlaSerAsnSerGly-----SerGlyProAspArgGluArgGlu 169
Db 3062 GATGAAAGCG 3012

QY 170 ThrThrProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAlaGly 189
Db 3011 ACCGATCCCGGTGGATCGACCGCGCGGACATCGGGGAC-----CGCGCG 2967

QY 190 HisLys-ThrGly-----ProSerLeuProAlaAlaThrProAlaAlaGluLeuIleVa 207
Db 2966 CACCGTCAGCGGTGGTCTCGCCCGCGCGCTCGATCCGGACCGCGCGCGCGCGCGCGCG 2907

QY 207 1-----ProProAlaHisGluIleGlnGluPhePheAlaAlaAlaGluAlaAlaGl 224
Db 2906 GACCAGGCG 2847

QY 224 nAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyValProLeuAspAl 244
Db 2846 CAGC-----GCGGCTCGTACTCTGTG----- 2825

QY 244 aGlyGlyArgPheGluTrpAlaProValValSer 255
Db 2824 ----GTGAAGACCCAGTGGCGCGCGCGTGCACACAGC 2795

RESULT 10
US-08-955-957A-6/c
Sequence 6, Application US/089555957A
Patent No. 6312920
GENERAL INFORMATION:
APPLICANT: DeHoff, Bradley S.
APPLICANT: Rostek Jr., Paul R.
TITLE OF INVENTION: SAM Operon
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/955,957A
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: P-10162
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 4848 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-955-957A-6

Alignment Scores:
Pred. No.: 0.316 Length: 4848
Score: 129.50 Matches: 65
Percent Similarity: 36.11% Conservative: 26
Best Local Similarity: 25.79% Mismatches: 91
Query Match: 9.93% Indels: 70
DB: 4 Gaps: 12

US-09-993-808B-2 (1-256) x US-08-955-957A-6 (1-4848)

QY 26 GlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyValAlaLysValAlaPro 45
Db 3404 GGGGAAGAACTCGTGGAGAGAGAGCGCGCTGCCGAGGATCTCGCGAGGTTGG 3345

QY 46 ArgArgLysArgAlaProAlaGlyGluProAlaAlaValSerAlaGlyGlyAspGly 65
Db 3344 TCCTACCGGCAGCCTCCCGCGGTCCGCCCGCGCGCGCGCGCGCGCGCGCGCGCAT 3285

QY 66 GlySerCysTyrIle-----His 71
Db 3284 GGCCCGTCCGGCGGTCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCAT 3225

QY 72 LeuArgSerArgMetLeuPheMetAlaProGlnProGlnProSerValAspSerVal 91
Db 3224 GTGCGG-----CGGCCAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3186

; SEQ ID NO 7
; LENGTH: 1143
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-457-568-7

Alignment Scores:

Pred. No.: 0.0983 Length: 1143
Score: 125.50 Matches: 80
Percent Similarity: 37.01% Conservative: 24
Best Local Similarity: 28.47% Mismatches: 88
Query Match: 9.62% Indels: 90
DB: 4 Gaps: 16

US-09-993-808B-2 (1-256) x US-09-457-568-7 (1-1143)

```
QY 9 ArgGlyAlaAlaGlyAlaGlu-----ValAlaAlaValGluValThrGlnValValGly 26
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 105 AGAGGAGGTGCGGGCGCTGCTGGAGGGCGGGCGCTGCCCAACGACCGGA-ATAGTTACG 163

QY 27 ValArgThrArgSerArgSer-----AlaAla-AlaThrGlyGlyValAlaLysVa 43
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 164 GTCGGAGCCGATCCAGGTTCATGATGATGGGCGAGCGCCGAGTGGCGGAGTGCT---GC 220

QY 43 lAlaProArgArg-----Ly 48
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 221 TGCTCCACGGCGGAGGCCCAACTGCGCGACCCCGCCACTCTACCCGACCCGTGCACG 280

QY 48 sArgAlaProAlaGlyGluProAlaAlaAlaValSerAlaGlyGlyAspGlySerCy 68
   :|||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 281 ACGTGCGCGGAGGGCTTCCT-----GGACACGCTGGTGGTGCT-- 320

QY 68 sTyrIleHisLeuArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerVa 88
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 321 -----GCACCGGGCGG----- 332

QY 88 lAspSerValProThrProValGluAlaAlaAspGlyAlaAlaGlyGln----- 104
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 333 -----GGCGCGGTGGACGTGCGCGTCAAACTGCGAGTCTAAACGGCGCCCTA 580

QY 105 -----GlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThrAl 120
   :|||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 359 GGGCGCTGCTGCCGTGGACCTGCTGAGGAGTGGGCGCATCGCGATGCGCGGTACC 418

QY 120 aSerSerValAsnLeuGlyLeuGlyGln-ArgGlySerHisThrCysArgSerTyrA 140
   :|||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 281 ACGTGCGCGGAGGGCTTCCT-----GGACACGCTGGTGGTGCT-- 320

QY 140 sAlaAlaGluAlaGlyAspHisValLeuValAspValSerAlaAlaSer---AsnS 159
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 467 ATGCCCGGAAGTCCCTCAGAC-----ATCCCCGATGGTGGCGGGGCGAGGGGTG 520

QY 159 erGlySerGlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyG 179
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 521 GCGGTCCGGGGGTGGATCCGTCGAGTCAAACTGCGAGTCTAAACGGCGCCCTA 580

QY 179 lAlaLeuSerAspLeuSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaA 199
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 581 GCCTGGAGCGGATGGACCGCAGGCGGCGGAGCAC-----CCCAAGC 622

QY 199 laThrProAlaAlaGluLeuIleValProProAlaHisGlu-----IleG 214
   :|||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 623 CCTCGGCTGCAGGAACCTCTTCGCGCCCGGTGGACCAAGAGTAAACCCGGGACTTGG 682

QY 214 lGluPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnP 234
   :|||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 683 AGAGGACTGCAGAGACATGGAAGAGCGGAGCCAGCGC-----AAGTGAATT 730

QY 234 heAspPheValArgGlyValProLeuAspAlaGlyArgPheGluTrpAlaProVal 253
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 731 TCGATTTTCAGAAATCAAAACCCCTAGAG-----GGCAAGTACGAGTGGCAAGAGGTG 783
```

RESULT 13

US-09-457-646-7
; Sequence 7, Application US/09457646
; Patent No. 6420345
; GENERAL INFORMATION:
; APPLICANT: Patel, Salil D
; APPLICANT: McArthur, James G
; APPLICANT: Gyuris, Jenio

; TITLE OF INVENTION: Methods and Reagents for Inhibiting Proliferation of
; TITLE OF INVENTION: Smooth Muscle Cells
; FILE REFERENCE: 106482.287
; CURRENT APPLICATION NUMBER: US/09/457,646
; CURRENT FILING DATE: 1999-12-09
; EARLIER APPLICATION NUMBER: 60/122,974
; EARLIER FILING DATE: 1999-03-01
; EARLIER APPLICATION NUMBER: 60/163,682
; EARLIER FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7

; LENGTH: 1143

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-457-646-7

Alignment Scores:

Pred. No.: 0.0983 Length: 1143
Score: 125.50 Matches: 80
Percent Similarity: 37.01% Conservative: 24
Best Local Similarity: 28.47% Mismatches: 88
Query Match: 9.62% Indels: 90
DB: 4 Gaps: 16

US-09-993-808B-2 (1-256) x US-09-457-646-7 (1-1143)

```
QY 9 ArgGlyAlaAlaGlyAlaGlu-----ValAlaAlaValGluValThrGlnValValGly 26
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 105 AGAGGAGGTGCGGGCGCTGCTGGAGGGCGGGCGCTGCCCAACGACCGGA-ATAGTTACG 163

QY 27 ValArgThrArgSerArgSer-----AlaAla-AlaThrGlyGlyValAlaLysVa 43
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 164 GTCGGAGCCGATCCAGGTTCATGATGATGGGCGAGCGCCGAGTGGCGGAGTGCT---GC 220

QY 43 lAlaProArgArg-----Ly 48
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 221 TGCTCCACGGCGGAGGCCCAACTGCGCGCGACCCCGCCACTCTACCCGACCCGTGCACG 280

QY 48 sArgAlaProAlaGlyGluProAlaAlaAlaValSerAlaGlyGlyAspGlySerCy 68
   :|||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 281 ACGTGCGCGGAGGGCTTCCT-----GGACACGCTGGTGGTGCT-- 320

QY 68 sTyrIleHisLeuArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerVa 88
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 321 -----GCACCGGGCGG----- 332

QY 88 lAspSerValProThrProValGluAlaAlaAspGlyAlaAlaGlyGln----- 104
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 333 -----GGCGCGGTGGACGTGCGCGTCAAACTGCGAGTCTAAACGGCGCCCTA 580

QY 105 -----GlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThrAl 120
   :|||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 359 GGGCGCTGCTGCCGTGGACCTGCTGAGGAGTGGGCGCATCGCGATGCGCGGTACC 418

QY 120 aSerSerValAsnLeuGlyLeuGlyGln-ArgGlySerHisThrCysArgSerTyrA 140
   :|||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 281 ACGTGCGCGGAGGGCTTCCT-----GGACACGCTGGTGGTGCT-- 320

QY 140 sAlaAlaGluAlaGlyAspHisValLeuValAspValSerAlaAlaSer---AsnS 159
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 467 ATGCCCGGAAGTCCCTCAGAC-----ATCCCCGATGGTGGCGGGGCGAGGGGTG 520

QY 159 erGlySerGlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyG 179
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 521 GCGGTCCGGGGGTGGATCCGTCGAGTCAAACTGCGAGTCTAAACGGCGCCCTA 580
```


QY 179 luleuSerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaA 199
Db 581 GCCTGGAGCGGATGGACGCCAGCGGAGCAC-----CCCAAGC 622
QY 199 laThrProAlaAlaGluLeuValProProAlaHisGlu-----IleG 214
Db 623 CCTCGGCTGCAGGAACCTCTTCGGCCCGGTGGACCAAGAGATTAAACCCGGGACTTG 682
QY 214 lngluPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnP 234
Db 683 AGAAGCACTGCAGAGACATGGAAGAGCGGAGCCAGCGC-----AAGTGGGAATT 730
QY 234 heAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGluTyrAlaProVal 253
Db 731 TCGATTTCAGATCAAAACCCCTAGAG-----GGCAAGTACGAGTGGCAAGAGGTG 783

RESULT 14
US-08-902-572-5
; Sequence 5, Application US/08902572
; Patent No. 6495526
; GENERAL INFORMATION:
; APPLICANT: Gyuris, Jeni
; APPLICANT: Lamphere, Lou
; APPLICANT: Beach, David H.
; TITLE OF INVENTION: INHIBITORS OF CELL-CYCLE PROGRESSION AND
; TITLE OF INVENTION: RELATED THERETO
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/902,572
; FILING DATE: 29-JUL-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-069.03
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1143 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1140
US-08-902-572-5

Alignment Scores:
Pred. No.: 0.0983 Length: 1143
Score: 125.50 Matches: 80
Percent Similarity: 37.01% Conservative: 24
Best Local Similarity: 28.47% Mismatches: 88
Query Match: 9.62% Indels: 90
DB: 4 Gaps: 16

US-09-993-808B-2 (1-256) x US-08-902-572-5 (1-1143)

QY 9 ArgGlyAlaAlaGlyAlaGlu-----ValAlaAlaValGluValThrGlnValValGly 26
Db 105 AGAGGAGGTGCGGGCGCTGCTGGAGCGGGGGCTGCCAACGACCGA-ATAGTTACG 163
QY 27 ValArgThrArgSerArgSer-----AlaAla-AlaThrGlyGlyValAlaLysVa 43
Db 164 GTCCGAGGCCGATCCAGGTCTATGATGATGGCAGCGCCGAGTGGCGGAGTGTCT--GC 220
QY 43 lAlaProArgArg-----ly 48
Db 221 TGTCCACGGCGGAGCCCAACTGCGCCGACCCCGCCACTCTCACCCGACCCGTGCACG 280
QY 48 sArgAlaProAlaGlyGluProAlaAlaAlaValSerAlaGlyGlyAspGlyGlySerCy 68
Db 281 ACCTGCCCGGGAGGGCTTCT-----GGACACGCTGGTGTGTCT-- 320
QY 68 sTyrIleHisLeuArgSerArgMetLeuPheMetAlaProGlnProGlnProSerVa 88
Db 321 -----GCACCGGCGCG----- 332
QY 88 lAspSerValProThrProValGluAlaAlaAspGlyAlaAlaGlyGln----- 104
Db 333 -----GGCGCGCTGGACGTGCGCGATGCGCT 358
QY 105 -----GlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThrAl 120
Db 359 GGGCGCGTCTGCCGTGGACCTGCTGAGGAGTGGCCATCGCGATGTCGACGGTACC 418
QY 120 aSerSerValAsnLeuGlyGlyGln-ArgGlySerHisThrCysArgSerTyrA 140
Db 419 TGCAGC-----GGCTGCGGGGGCACCAGAGGAGTAACCATGCCCGC---ATAG 466
QY 140 sAlaAlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSer---Asn 159
Db 467 ATGCGCGGAAGGTCCCTCAGAC-----ATCCCCGATGTTGGCGGGGCGGCGGTG 520
QY 159 erGlySerGlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyG 179
Db 521 GCGGTCCGGCGGGGTGGATCCGTCGAGTCAAACGTGCGAGTGTCTAACGGCGCCCTA 580
QY 179 luLeuSerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAla 199
Db 581 GCCTGAGCGGATGGACGCCAGCGGCGGAGCAC-----CCCAAGC 622
QY 199 laThrProAlaAlaGluLeuValProProAlaHisGlu-----IleG 214
Db 623 CCTCGGCTGCAGGAACCTCTTCGGCCCGGTGGACCAAGAGATTAAACCCGGGACTTG 682
QY 214 lngluPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnP 234
Db 683 AGAAGCACTGCAGAGACATGGAAGAGCGGAGCCAGCGC-----AAGTGGGAATT 730
QY 234 heAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGluTyrAlaProVal 253
Db 731 TCGATTTCAGATCAAAACCCCTAGAG-----GGCAAGTACGAGTGGCAAGAGGTG 783

RESULT 15
US-09-516-065-7
; Sequence 7, Application US/09516065
; Patent No. 6521602
; GENERAL INFORMATION:
; APPLICANT: Patel, Salil D
; APPLICANT: McArthur, James G
; APPLICANT: Gyuris, Jeni
; APPLICANT: Mendez, Michel
; APPLICANT: Finer, Mitchel
; TITLE OF INVENTION: Anti-Neoplastic Compositions and Uses Thereof
; FILE REFERENCE: Cell 406; 106482.692
; CURRENT APPLICATION NUMBER: US/09/516,065
; CURRENT FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/128,515
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/128,271

;
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: 60/122,974
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1143
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-516-065-7

Alignment Scores:
Pred. No.: 0.0983 Length: 1143
Score: 125.50 Matches: 80
Percent Similarity: 37.01% Conservative: 24
Best Local Similarity: 28.47% Mismatches: 88
Query Match: 9.62% Indels: 90
DB: 4 Gaps: 16

US-09-993-808B-2 (1-256) x US-09-516-065-7 (1-1143)

QY 9 ArgGlyAlaAlaGlyAlaGlu-----ValAlaAlaValGluValThrGlnValValGly 26
Db 105 AGAGGAGTGGGGCGCTGCTGGAGGGGGGGCGCTGCCAACGCCACCGA-ATAGTTACG 163
QY 27 ValArgThrArgSerArgSer-----AlaAla-AlaThrGlyGlyValAlaLysVa 43
Db 164 GTGGAGCGCGATCCAGGTTCATGATGATGGCAGCGCCCGAGTGGCGGAGTGCT---GC 220
QY 43 lAlaProArgArg-----Ly 48
Db 221 TGCTCCACGGCGGAGGCCCAACTGGCGGACCCCGCCACTCTCACCGACCCCGTGCAAG 280
QY 48 sArgAlaProAlaGlyGluProAlaAlaAlaValSerAlaGlyGlyAspGlyGlySerCy 68
Db 281 ACGCTGCCCGGGAGGGCTTCCT-----GGACACGCTGGTGGTGCT-- 320
QY 68 sTyrIleHisLeuArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerVa 88
Db 321 -----GCACCGGGCGG----- 332
QY 88 lAspSerValProThrProValGluAlaAlaAspGlyAlaAlaGlyGln----- 104
Db 333 -----GGCGGGCTGGACGTGCGGATGCCT 358
QY 105 -----GlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerThrAl 120
Db 359 GGGCGCTGCTGCCGTGGACCTGGTGAGGAGCTGGGCCATCGCGATGTCGCACGGTACC 418
QY 120 aSerSerValAsnLeuGlyLeuGlyGlyGln-ArgGlySerHisThrCysArgSerTyrA 140
Db 419 TGGCGG-----GGCTGGGGGGGCACAGAGGAGTAAACCATGCCCGC---ATAG 466
QY 140 spAlaAlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSer---AsnS 159
Db 467 ATGCCCGGAAGGTCCCTCAGAC-----ATCCCCGATGGTGGCGGGGCGAGCGGGGGTG 520
QY 159 erGlySerGlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyG 179
Db 521 GCGGTTCGGGGGGGGTGGATCCGTCGAGTCAAAACGTGCGAGTGTCTAACGGGGCGCCCTA 580
QY 179 luleuSerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaA 199
Db 581 GCCTGGAGCGGATGGACGCCAGCGAGGCGGAGCAC-----CCCCAGC 622
QY 199 laThrProAlaAlaGluLeuIleValProProAlaHisGlu-----IleG 214
Db 623 CCTCGGCTGCAGGAACCTCTTCGGCCCGGTGGACCAACGAGAGTTAACCCGGGACTTGG 682
QY 214 lnGluPhePheAlaAlaAlaGluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnP 234
Db 683 AGAAGCACTGCAGAGACATGGAGAGGCGGAGCCAGCGC-----AAGTGGATT 730

QY 234 heAspPheValArgGlyValProLeuAspAlaGlyGlyArgPheGluTrpAlaProVal 253
Db 731 TCGATTTTCAGATCATCAAAACCCCTAGAG-----GGCAAGTACGAGTGGCAAGAGGTG 783

Search completed: October 2, 2004, 12:44:01
Job time : 3166 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2004, 11:11:40 ; Search time 447 Seconds
(without alignments)
2903.376 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 1304
Sequence: 1 MGKYMRCRGAAGAEVAE.....VRGVPLDAGGRFEWAPVWSI 256

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3340653 seqs, 2534783454 residues

Total number of hits satisfying chosen parameters: 6681306

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09933808/runat_01102004_171704_13115/app_query.fasta_1.455
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09933808@cgn_1_1_723@runat_01102004_171704_13115
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US09_PUB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query | Score | Match | Length | DB | ID | Description |
|------------|-------|-------|-------|--------|----|----|-------------|
|------------|-------|-------|-------|--------|----|----|-------------|

RESULT 1
US-09-993-308-1
; Sequence 1, Application US/099933308
; Patent No. US20020159435A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William J.
; APPLICANT: Lowe, Keith S.
; APPLICANT: Larkins, Brian A.
; APPLICANT: Dilkes, Brian R.
; APPLICANT: Sun, Yuejin
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,308
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 60/246,349
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1372
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:

ALIGNMENTS

| | | | | | | | |
|---|----|--------|-------|------|----|----------------------|-------------------|
| | 1 | 1304 | 100.0 | 1372 | 9 | US-09-993-308-1 | Sequence 1, Appli |
| | 2 | 1304 | 100.0 | 1372 | 11 | US-09-993-808B-1 | Sequence 1, Appli |
| | 3 | 1217.5 | 93.4 | 1111 | 13 | US-10-425-114-34947 | Sequence 34947, A |
| c | 4 | 1059.5 | 81.2 | 985 | 13 | US-10-425-114-18379 | Sequence 18379, A |
| | 5 | 669 | 51.3 | 1242 | 16 | US-10-333-006-9 | Sequence 9, Appli |
| c | 6 | 669 | 51.3 | 1671 | 17 | US-10-437-963-92632 | Sequence 92632, A |
| c | 7 | 668.5 | 51.3 | 834 | 13 | US-10-425-114-18085 | Sequence 18085, A |
| | 8 | 664 | 50.9 | 841 | 9 | US-09-993-308-5 | Sequence 5, Appli |
| | 9 | 664 | 50.9 | 841 | 11 | US-09-993-808B-5 | Sequence 5, Appli |
| | 10 | 512 | 39.3 | 510 | 17 | US-10-767-701-24179 | Sequence 24179, A |
| | 11 | 399.5 | 30.6 | 568 | 16 | US-10-333-006-5 | Sequence 5, Appli |
| | 12 | 315.5 | 24.2 | 1065 | 17 | US-10-437-963-19949 | Sequence 19949, A |
| c | 13 | 314 | 24.1 | 1398 | 17 | US-10-437-963-19960 | Sequence 19960, A |
| | 14 | 279 | 21.4 | 681 | 16 | US-10-333-006-54 | Sequence 54, Appl |
| | 15 | 242 | 18.6 | 1073 | 16 | US-10-333-006-56 | Sequence 56, Appl |
| | 16 | 238 | 18.3 | 906 | 17 | US-10-437-963-18387 | Sequence 18387, A |
| | 17 | 238 | 18.3 | 1183 | 17 | US-10-437-963-96091 | Sequence 96091, A |
| | 18 | 235 | 18.0 | 1067 | 16 | US-10-333-006-43 | Sequence 43, Appl |
| | 19 | 231 | 17.7 | 1110 | 13 | US-10-425-114-21130 | Sequence 21130, A |
| | 20 | 218 | 16.7 | 1360 | 13 | US-10-424-599-92958 | Sequence 92958, A |
| | 21 | 215.5 | 16.5 | 797 | 17 | US-10-451-139-20 | Sequence 20, Appl |
| | 22 | 215 | 16.5 | 1112 | 13 | US-10-425-114-25318 | Sequence 25318, A |
| | 23 | 213 | 16.3 | 898 | 13 | US-10-425-114-22184 | Sequence 22184, A |
| | 24 | 212 | 16.3 | 794 | 17 | US-10-437-963-25722 | Sequence 25722, A |
| | 25 | 205.5 | 15.8 | 1261 | 13 | US-10-424-599-67348 | Sequence 67348, A |
| | 26 | 200 | 15.3 | 1089 | 9 | US-09-993-308-3 | Sequence 3, Appli |
| | 27 | 200 | 15.3 | 1089 | 11 | US-09-993-808B-3 | Sequence 3, Appli |
| | 28 | 199.5 | 15.3 | 804 | 9 | US-09-733-507-15 | Sequence 15, Appl |
| | 29 | 199.5 | 15.3 | 804 | 17 | US-10-451-139-14 | Sequence 14, Appl |
| | 30 | 190.5 | 14.6 | 875 | 12 | US-10-688-291-3 | Sequence 3, Appli |
| | 31 | 186 | 14.3 | 997 | 13 | US-10-425-114-33848 | Sequence 33848, A |
| | 32 | 178 | 13.7 | 1290 | 17 | US-10-451-139-16 | Sequence 16, Appl |
| | 33 | 165 | 12.7 | 756 | 17 | US-10-767-701-5691 | Sequence 5691, Ap |
| | 34 | 163 | 12.5 | 1237 | 13 | US-10-424-599-40086 | Sequence 40086, A |
| c | 35 | 152 | 11.7 | 4887 | 17 | US-10-437-963-34324 | Sequence 34324, A |
| | 36 | 151.5 | 11.6 | 493 | 16 | US-10-333-006-1 | Sequence 1, Appli |
| | 37 | 150.5 | 11.5 | 1040 | 13 | US-10-424-599-69339 | Sequence 69339, A |
| | 38 | 149 | 11.4 | 984 | 9 | US-09-938-842A-27 | Sequence 27, Appl |
| | 39 | 149 | 11.4 | 984 | 11 | US-09-938-842A-27 | Sequence 27, Appl |
| | 40 | 145 | 11.1 | 533 | 9 | US-09-733-507-9 | Sequence 9, Appli |
| | 41 | 145 | 11.1 | 533 | 17 | US-10-451-139-9 | Sequence 9, Appli |
| c | 42 | 143.5 | 11.0 | 2328 | 15 | US-10-114-153-57 | Sequence 57, Appl |
| c | 43 | 142 | 10.9 | 580 | 13 | US-10-425-114-16891 | Sequence 16891, A |
| c | 44 | 141.5 | 10.9 | 4281 | 13 | US-10-282-122A-13341 | Sequence 13341, A |
| c | 45 | 140.5 | 10.8 | 2787 | 17 | US-10-437-963-11239 | Sequence 11239, A |

; NAME/KEY: CDS
 ; LOCATION: (134)..(904)
 ; OTHER INFORMATION:
 US-09-993-308-1

| Alignment Scores: | |
|------------------------|----------|
| Pred. No.: | 2.2e-118 |
| Score: | 1304.00 |
| Percent Similarity: | 100.00% |
| Best Local Similarity: | 100.00% |
| Query Match: | 100.00% |
| DB: | 9 |
| Length: | 1372 |
| Matches: | 256 |
| Conservative: | 0 |
| Mismatches: | 0 |
| Indels: | 0 |
| Gaps: | 0 |

US-09-993-808B-2 (1-256) x US-09-993-308-1 (1-1372)

| | | | |
|----|-----|---|-----|
| QY | 1 | MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu | 20 |
| DB | 134 | ATGGGAAGTACATCGCAAGTGCAGGGCGCGCAGGCGCGGAGGTCGCGCGGT | 193 |
| QY | 21 | ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal | 40 |
| DB | 194 | GTTACGCAGGTCGTCGGCGTCCGGACGAGGTCAGGTCGCGGGCGGACCGCGCGGTGTC | 253 |
| QY | 41 | AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaAlaValSer | 60 |
| DB | 254 | GCGAAGGTCGCCCCGAGGAGGAAGAGGGCGCGGGGGAGCCCTGCTGCCGCGTGAGC | 313 |
| QY | 61 | AlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAla | 80 |
| DB | 314 | GCTGTTGGGACGGCGGAAGCTGCTACATCCACCTGCGTAGCCGCATGCTGTTTCATGGCA | 373 |
| QY | 81 | ProProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAspGly | 100 |
| DB | 374 | CCGCCTCAGCCGCGAGCCGTCGGTTGACTCGGTTCCGACCCCGGTGGAGGTCGCTGATGGC | 433 |
| QY | 101 | AlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThrAla | 120 |
| DB | 434 | GCTGCAGGACAGCAGGGCGCGGCTCGCGGCCGGGCTCTCGCGTTGCTCCAGCACGGCG | 493 |
| QY | 121 | SerSerValAsnLeuGlyLeuGlyGlyGlnArgGlySerHisThrCysArgSerTyrAsp | 140 |
| DB | 494 | TCGTCGGTGAACCTTGGGCTTGGGGGTTCAGCGCGGAGCCACACCTGCCGCTCCTACGAC | 553 |
| QY | 141 | AlaAlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGly | 160 |
| DB | 554 | GCTGCAGAGGCTGGCGGGATCACGTCCTGGTGGATGTCCTCGCGGCGAGCAACTCCGGG | 613 |
| QY | 161 | SerGlyProAspArgGluArgArgGluThrThrProSerSerArgAlaHisGlyGluLeu | 180 |
| DB | 614 | AGCGGCCAGACCCGAGAGGCGAGAGACGACGCCATCGAGCCGGCGCACCGCGGAGCTC | 673 |
| QY | 181 | SerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThr | 200 |
| DB | 674 | AGCGATCTGGAGTCGGATCTGGCGGGGCACAAGACTGGCCCCGTCGCTACCGGCGGCAACG | 733 |
| QY | 201 | ProAlaAlaGluLeuIleValProProAlaHisGluIleGlnGluPhePheAlaAlaAla | 220 |
| DB | 734 | CCGGCTGGCGAGCTGATCGTCCGCCACGACACGAGATCCAGGAGTTCTTCGCCGCCGCC | 793 |
| QY | 221 | GluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyVal | 240 |
| DB | 794 | GAGCGGCCACAGGCCAAGCGCTTTGCTTCCAAGTACAACTTCGACTTCGTCCGCGGCGTG | 853 |
| QY | 241 | ProLeuAspAlaGlyGlyArgPheGluTrpAlaProValValSerIle | 256 |
| DB | 854 | CCCTCTGACGCCGCGGCGGTTTCGAGTGGCGCGCGGTGGTTCAGCATCTC | 901 |

RESULT 2

US-09-993-808B-1

03-09-2025 08:05:11
; Sequence 1, Application US/09993808B
; Publication No. US20040003433A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William

```

; APPLICANT: Lowe, Keith
; APPLICANT: Sun, Yuejin
; APPLICANT: Dilkes, Brian
; APPLICANT: Larkins, Brian
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides,
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,808B
; CURRENT FILING DATE: 2001-11-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1372
; TYPE: DNA
; ORGANISM: zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (134)...(902)
US-09-993-808B-1

```

| | | |
|------------------------|---------------|------|
| Alignment Scores: | | |
| Pred. No.: | 2.20e-118 | 1372 |
| Score: | 1304.00 | 256 |
| Percent Similarity: | 100.00% | 0 |
| Best Local Similarity: | 100.00% | 0 |
| Query Match: | 100.00% | 0 |
| DR. | 11 | 0 |
| | | DR. |
| | Length: | |
| | Matches: | |
| | Conservative: | |
| | Mismatches: | |
| | Indels: | |
| | Gaps: | |

US-09-993-808B-2 (1-256) x US-09-993-808B-1 (1-1372)

| | | | |
|----|-----|---|-----|
| QY | 1 | MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu | 20 |
| DB | 134 | ATGGGAAGTACATGCGCAAGTGCAGGGGCGCCGACAGGCGGAGGTGCGCCCGCTCGAG | 193 |
| QY | 21 | ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal | 40 |
| DB | 194 | GTTACGCAGGTGCTCGGCGTCCGACGAGTCCAGTCCGCGCGCGACCGCGGTGTC | 253 |
| QY | 41 | AlaLysValAlaProArgLysArgAlaProAlaGlyGluProAlaAlaAlaValSer | 60 |
| DB | 254 | GCGAAGTGCCTCCGAGGAGGAGAGGGCGCCGCGGGGAGCCGTGTCGCCCGCTGAGC | 313 |
| QY | 61 | AlaGlyGlyAspGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAla | 80 |
| DB | 314 | GCTGTTGGGACGCGCGAAGCTGCTACATCCACCTGCGTAGCCGCATGCTGTTCATGGCA | 373 |
| QY | 81 | ProProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAspGly | 100 |
| DB | 374 | CCGCCTCAGCCGACGCCGTGCGTTGACTCGGTTCCGACCCCGTGGAGGCTGCTGATGGC | 433 |
| QY | 101 | AlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThrAla | 120 |
| DB | 434 | GCTGCAGGACAGCAGGGCGCGGCTTCGCGCCCGGCTCTCGCGTTGCTCCAGCACGGCG | 493 |
| QY | 121 | SerSerValAsnLeuGlyLeuGlyGlyGlnArgGlySerHisThrCysArgSerTyrAsp | 140 |
| DB | 494 | TCGTTCGGTGAACCTTGGGCTTGGGGGTGACGCGGGAGCCACACCTGCCGTCCTACGAC | 553 |
| QY | 141 | AlaAlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGly | 160 |
| DB | 554 | GCTGACAGGCTGGCGGGATCACGTCCTGGTGGATGTCTCGGGCGGAGCAACTCCGGG | 613 |
| QY | 161 | SerGlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyGluLeu | 180 |
| DB | 614 | AGCGGCCAGACCCGAGAGGCGAGACGACGCCATCGAGCCGGCGCACGGCGAGCTC | 673 |
| QY | 181 | SerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThr | 200 |
| DB | 674 | AGCATCTGGAGTCGGATCTGGCGGGGCACAAGACTGGCCCGTCGCTACCGGCGGCAACG | 733 |
| QY | 201 | ProAlaAlaGluLeuIleValProProAlaHisGluIleGlnGluPhePheAlaAlaAla | 220 |
| DB | 734 | CCGGCTGCGGAGCTGATCGTGCCTGCCGACACACGAGATCCAGGAGTTCTTCGCCCGCGCC | 793 |

| | | | |
|----|-----|--|-----|
| Qy | 201 | ProAlaAlaGluLeuIleValProProAlaHisGluIleGlnGluPhePheAlaAlaAla | 220 |
| | | | |
| Db | 734 | CCGGTTCGGAGCTGATCGTGCCGCCAGCACCGAGATCCAGGAGTTCTTCGCCGCCGCC | 793 |

QY 221 GluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyVal 240
Db 794 GAGCGGCCAGGCCAAGCGCTTTGCTTCCAAAGTACAACTTCGACTTCGTCCGCGCGTG 853
QY 241 ProLeuAspAlaGlyGlyArgPheGluTrpAlaProValValSerIle 256
Db 854 CCCCCTGACGCCGCGCGGCTTCGAGTGGCGCGCGGTGGTCAGCATC 901
RESULT 3
US-10-425-114-34947
; Sequence 34947, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 34947
; LENGTH: 1111
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17355F05_FLI
US-10-425-114-34947
Alignment Scores:
Pred. No.: 5.18e-110 Length: 1111
Score: 1217.50 Matches: 245
Percent Similarity: 96.50% Conservative: 3
Best Local Similarity: 95.33% Mismatches: 4
Query Match: 93.37% Indels: 5
DB: 13 Gaps: 3
US-09-993-808B-2 (1-256) x US-10-425-114-34947 (1-1111)
QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
Db 32 ATGGGGAAGTACATCGCAAGTGCAGGGGGGCGCGAGGCGCGAGGTGCGCCCGCTCGAG 91
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal 40
Db 92 GTTACGAGGTGCTGGCGTCCGACAGAGGTCCAGTCCGCGCGCGAGCCGCGGTGTC 151
QY 41 AlaLys--ValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaVal 59
Db 152 GCGAAGGTGCTCGCCCCGAGGAGGAGGCGCGCGCGCGGAGCCTGCTGCCTCCGTG 211
QY 60 SerAlaGlyGlyAspGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMet 79
Db 212 GCGCTGGTGGGACGCGGGAAGCTGTACATCCACCTGCTAGCGCATGCTGTTCATG 271
QY 80 AlaProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAsp 99
Db 272 GCACCGCTCAGCCGACGCCG-----CCGTCGGTTCCGACCCCGCGGAGGTGCTGAT 325
QY 100 GlyAlaAlaGlyGlnGlnAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThr 119
Db 326 GCGGTGCAGGACAGCAGGCGCGCGCTCGCGCGCGGCTCTCGCGTTCAGCAGCAG 385
QY 120 AlaSerSerValAsnLeuGlyLeuGlyGlnArgGlySerHisThrCysArgSerTyr 139
Db 386 GCGTCGTGGTGCAC-----GTGGGGGTTCAGCGCGGAGCACACCTGCCGCTCCGAC 439
QY 140 AspAlaAlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSer 159

Db 440 GACGCTGCAGAGGCTGGCGGGATCACGTCCTGGTGGATGTCTCGCGCGCGAGCAACTCC 499
QY 160 GlySerGlyProAspArgGluArgArgGluThrThrProSerSerArgAlaHisGlyGlu 179
Db 500 GGGAGCGGCCAGACCGGAGAGGCGAGAGACGACGCGCATCGAGCGGCGCGCAGCGAG 559
QY 180 LeuSerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaAla 199
Db 560 CTCAGCGATCTGGAGTCGATCTGGCGGGGCACAAGACTGGCCCGCTCGCTACCGCGGCA 619
QY 200 ThrProAlaAlaGluLeuIleValProProAlaHisGluIleGlnGlnPheAlaAla 219
Db 620 ACGCCGCTGGGAGTGTGTCGCCCGCAGCACACGAGATCCAGGAGTTCTTCGCCGCC 679
QY 220 AlaGluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGly 239
Db 680 GCCGAGCGGCCAGGCCAAGCGCTTGTCTTCCAAAGTACAACTTCGACTTCGTCCGTGGC 739
QY 240 ValProLeuAspAlaGlyGlyArgPheGluTrpAlaProValValSerIle 256
Db 740 GTGCCCTCGACGCCGCGCGCGGTTCGAGTGGCGCGCGGTGCTGAGCATC 790
RESULT 4
US-10-425-114-18379/c
; Sequence 18379, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 18379
; LENGTH: 985
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LTB3079-036-F9_FLI
US-10-425-114-18379
Alignment Scores:
Pred. No.: 1.34e-94 Length: 985
Score: 1059.50 Matches: 213
Percent Similarity: 96.00% Conservative: 3
Best Local Similarity: 94.67% Mismatches: 4
Query Match: 81.25% Indels: 5
DB: 13 Gaps: 3
US-09-993-808B-2 (1-256) x US-10-425-114-18379 (1-985)
QY 33 SerAlaAlaAlaThrGlyValAlaLys---ValAlaProArgArgLysArgAlaPro 51
Db 984 TCCGCGCGCGCGACCGCGCGTGTGCGAAGGTGTCGCCCGCAGGAGGAGCGCGCG 925
QY 52 AlaGlyGluProAlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHis 71
Db 924 GCGGGGAGCCTGCTGCCTCCGTGGCGCTGTTGGGACGCGGAGCTGTACATCCAC 865
QY 72 LeuArgSerArgMetLeuPheMetAlaProGlnProGlnProSerValAspSerVal 91
Db 864 CTGCGTAGCCGATGCTGTTCATGGCACCCTCAGCCGCGCGCGCGCGCGCGCGCG 811
QY 92 ProThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAla 111
Db 810 CCGACCCCGCGGAGGCTGCTGATGGCGCTGCAGGACAGCGCGCGCGCGCGCGCGCG 751


```

; NUMBER OF SEQ ID NOS: 2049666
; SEQ ID NO 92632
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID
US-10-437-963-92632

```

| | | |
|------------------------|----------|------------------|
| Alignment Scores: | | |
| Pred. No.: | 4.04e-56 | 1671 |
| Score: | 669.00 | 161 |
| Percent Similarity: | 67.42% | Conservative: 19 |
| Best Local Similarity: | 60.30% | Mismatches: 67 |
| Query Match: | 51.30% | Indels: 20 |
| DB: | 17 | Gaps: 10 |

US-09-993-808B-2 (1-256) x US-10-437-963-92632 (1-1671)

```

QY      1  MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu  20
      |||||
Db      1206  ATGGCAAGTACATGAGAAAGTTCAGGGGGGCCACGGGGGAGGAGTTGGCCCGCATGGAG  1147
      |||||
QY      21  ValThrGlnValValGlyValArgThrArgSerAlaAlaAlaThrGlyGly---  39
      |||||
Db      1146  GTCACGCAGGTGTTGGCGTCCGGACGAGGTCGAGGTCGGGCGGCGGCGGCGGCGGCGGCGG  1087

```

```

QY      40  -----ValAlaLysValAlaProArgArgLysArgAlaProAlaGlyGlu 54
Db      1086  ACGACGAAGGTGAAGCGCGCTCGGCGGCTCCACGAGGAGGAGGCGCTGCTGCCG 1027

QY      55  ProAlaAlaIvalSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSer 74
Db      1026  ACGGCGGTGCTGGGGACTACTCGCGGTGACGGCGGGAGCTGCTACCTCCAGCTGAGGAGC 967

```

[illegible]

| | | | |
|----|-----|--|-----|
| Qy | 95 | ValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSer | 114 |
| | | | |
| Db | 909 | GCGAGGCGGCG--GGTTCGGGAACGAGCGGCGGCATGCGGCGGTGGCCCTCTCG | 853 |

| | | | |
|----|-----|--|-----|
| Qy | 115 | ArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlnArgGlySerHis | 134 |
| | | | |
| Db | 852 | CGTTGCTCCAGCACGGCGTCGTCCTGGAC-----GCGGCGGCTCAGGACAGGAGCCTC | 799 |
| | | | |

QY 135 ThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuValAspValSer 154
Db 798 GCGTGCCGCTCC---GACGTCCGGAGGCAGGCAGCGAGCAT-----GTCCCGGAGGGC 748

Qy 155 AlaAlaSerAsnSerGlySerGlyProAspArgGluArgArgGluThrThrProSerSer 171
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 747 TCCGCGAGCGACTCGGCGAGCGGCCGTGACCGCGAGAGGAGAGAAACAATCCATCAAGC 688

Qy 175 ArgAlaHisGlyGluLeuSerAspLeuGluSerAspLeuAla---GlyHisLysThrGly 193
Db 687 TTTTCTCCCGGAGAGGTGAGCGATCTGGAGTCTGGCTGGAGGACAGAAAGCGCAGC 628

DQ
194 ProSerLeupro---AlaAlathProAlaAlaGluLeule-----Valpro 208

DB
627 CGTCCACTACCTTCTGCGGGCAACAGCCTCAGCACAGCCAGCCGGGCCGAAGATTCCG 568

QY 209 FFOALAHISGIUILEGINGIUPNEPNEAIAAIAAGIUAIAAIAAGIUAIALALYSARGPNE 222
Db 567 CCGGCCGCCGAGATCGAGGCGTTCTTCGCGGCGGCCGAGGAGGCTGAGGCCAAGCGCTTC 508

QY
229 ATATSEILYSIVLASNPNFNEASPPNEVALARGGLVAFPLEUASPALAGLYGARYGPNE 244
|||::|||
Db 507 GCCGCCAAGTACAACCTCGACGTCTCGGGCGTGCCCCCTCGACGCC---GGTCGGTTTC 451

| QY | Z#3 | 9TUU1P | PHATAFI | LOVAL | VALSEI | Z33 |
|----|-----|--------|---------|-------|--------|-----|
| | | | | | | |

Db 450 GAGTGGACTCCGGTGTGCAGC 430

RESULT 7

US-10-425-114-18085/c

; Sequence 18085, Application US/10425114

; Publication No. US2004003488A1

; GENERAL INFORMATION:

| | | | |
|------------------------|----------|---------------|-----|
| Alignment Scores: | | | |
| Pred. No.: | 2.16e-56 | Length: | 834 |
| Score: | 668.50 | Matches: | 174 |
| Percent Similarity: | 76.05% | Conservative: | 7 |
| Best Local Similarity: | 73.11% | Mismatches: | 30 |
| Query Match: | 51.27% | Indels: | 27 |
| DB: | 13 | Gaps: | 17 |

US-09-993-808B-2 (1-256) x US-10-425-114-18085 (1-834)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
Db 696 ATGGGGAAGTACATGGCAACGGCAGGGGGGGCCGGGGCGAGGGGGTGGCCGACATCCAG 637

| | | | | | | | | | | | | | | | | | | | | | |
|----|-----|------|-----|------|-----|-----|-----|-----|-----|-----|------|------|-----|-----|-----|-----|-----|-----|------|-----|----|
| Qy | 21 | Val | Thr | Gln | Val | Val | Gly | Val | Arg | Thr | Arg | Ser | Arg | Ser | Ala | Ala | Thr | --- | Gly | Gly | 39 |
| | | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | | |
| Db | 636 | GTCT | CGC | AGGT | CGT | CGG | CGT | CCG | GAC | GAG | GTCC | AGGT | CCG | CGG | CGG | GAC | CCG | CGG | CGGT | 577 | |

Qy 40 ValAlaLysValAla---ProArgArgLysArgAla-----ProAla----- 52
 ||||| ||||| ||||| ||||| |||||
Db 576 GTCGCGAAGGTGCTCCGCCGAGGAGGAAGAAGCGCTGCTGCCCGCCGGACGAGACG 517

QY 53 -----GlyGluProAlaAlaAlaValSerAlaGlyGlyAspGlyGlySer---CysTyr 69
 Db 516 GCCTCGGGGAGCCTGGTGCGTGGCGCTGGTGGTGGGACGGCGGAAGCTGCTGCTAC 457

Qy 70 IleHisLeuArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAsp 89
 |||||
 Db 456 ATCCACCTGCGGAGCCGCATGCTGTTTCATGGCAGCACCTCAGCAGCAACCGTCG----- 403

Qy 90 SerValProThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGly---AlaAla 108
:: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 402 GCGGCTCTACGCCGGTGGAGGTGCT---GGTCGGCA CAGCAGGGCGGGGTGGTGCGG 346

Qy 109 LeuAlaAlaGlyLeuSerArgCysSerSerThrAlaSerValAsnLeuGlyLeuGly
 Db 345 CTGCGGCTGGCCTCTCGCGTGTCTCCAGCAGCGCGTCTGTCGGTGGAC-----GTGCGG 292

QY
129 GYGInArg-----GlySerHisThrCysArgserTyrAspAla---AlaGluAla 148
|||:::
Db 291 GGCCAGCAGCCCGGAGCGGGAGCCACGCCTGCCGCTCCGACGCTGCGCCCTCGGAGGTT 232

[illegible]


```
QY 53 -----GlyGluProAlaAlaAlaValSerAlaGlyGlyAspGlyGlySer---CysTyr 69
Db 339 ACGTCGGGGAGCCTGGTCCCGTGGCGCTGGTGGTGGGGACGGCGAAGCTGCTGTAC 398
QY 70 IleHisLeuArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAsp 89
Db 399 ATCCACCTCGGAGCCGCGATGCTGTTTCATGGCAGCACCTCAGCAGCAACCGTCG----- 452
QY 90 SerValProThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGly---AlaAla 108
Db 453 GCGGCTCTGACGCGCGTGGAGGCTGCT---GGTGGCGCACARCAAGCGGGGTGGTGGCG 509
QY 109 LeuAlaAlaGlyLeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGly 128
Db 510 CTCGCGCTGGCCTCTCGCGTTGCTCCAGCACGGCGTCTCGTGGTGGAGTCTCGGAGTGC----- 560
QY 129 GlyGlnArgGlySerHisThrCysArgSerTyrAspAla---AlaGluAlaGlyGlyAsp 147
Db 561 -----GGGGCCACCGCTGCGCTCCGACGCTGCGCTCGCGCTCGGAGGTTGACGGGGAT 611
QY 148 HisValLeuValAspValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArg 167
Db 612 CACGTC---CCGGATGTCTGTCACCGCGAGCAACTCGGGGAGCGTCCCGACCGGAGAGG 668
QY 168 ArgGluThrThrPro---SerSerArgAlaHis---GlyGluLeuSerAspLeuGluSer 185
Db 669 AGAGAGACGACGCCATCGTCGAGCGGGCGCACGGCGGAGTCTGAGTCTGCGGATCG 728
QY 186 AspLeuAlaGly---HisLysThrGly---ProSerLeuProAla---AlaThrProAla 202
Db 729 GATCTGGTGGGGCGGAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 788
QY 203 AlaGluLeuIleValProProAlaHisGluIleGlnGluPhePheAlaAla 219
Db 789 GCGGAGCTGATCGTGCGCGCCAGCACAGGAGATCCAGGAATTCTTCGCGGCC 839

RESULT 10
US-10-767-701-24179
; Sequence 24179, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 24179
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 30165606
US-10-767-701-24179
```

```
Alignment Scores:
Pred. No.: 2.69e-41 Length: 510
Score: 512.00 Matches: 112
Percent Similarity: 85.82% Conservative: 3
Best Local Similarity: 83.58% Mismatches: 5
Query Match: 39.26% Indels: 14
DB: 17 Gaps: 5
```

US-09-993-808B-2 (1-256) x US-10-767-701-24179 (1-510)

```
QY 1 MetGlyLysTyrMetArg---LysCysArgGlyAlaAlaGlyAlaGluValAlaAlaVal 19
Db 121 ATGGGGAAGTACATGCGCAGCAAGTCAGGGGGCCCGGGGCGAGGAGTCCGCCGAGTC 180
```

```
QY 20 GluValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaThrGlyGly 39
Db 181 GAGGTCACGCAGGTTGTCCGCGTCAGGACGAGGTCAGGTCAGGTCGCGCGGCGACCGCGGT 240
QY 40 ValAlaLysValAlaProArgArgLysArgAla-----ProAla 52
Db 241 GTCGGAAGGTCGCTCCGAGGAGGAAGAGCGCTGACGCCCGCGCGAACGTGCCGCG 300
QY 53 GlyGluProAlaAlaAlaValSerAlaGlyGlyAspGlyGlySerCysTyrIleHisLeu 72
Db 301 GGGGAGCCT---GCCGCCGTGGCGCTGGCGGGGACGGCGGAGTGTACTACATCCACCTG 357
QY 73 ArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAspSerValPro 92
Db 358 CGAAGCCGCATGCTGTTTCATGGCACCGCCTCAGCCGCGACCG-----TCGGTTCCG 408
QY 93 ThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaGly 112
Db 409 ACGCCGCGGAGGCTGCT-----GGTGCAGGACAGCAGGGCGCGGCTCGTGGCTGGG 462
QY 113 LeuSerArgCysSerSerThrAlaSerSerValAsnLeuGly 126
Db 463 CTCTCGGTTGCTCCAGCACGGCGTCTCGTGGTGGACGTCCGG 504

RESULT 11
US-10-333-006-5
; Sequence 5, Application US/10333006
; Publication No. US20040019926A1
; GENERAL INFORMATION:
; APPLICANT: Frankard, Valerie Marie-No. US20040019926A1lle S.
; APPLICANT: Peres Bota, Adrian Marius
; APPLICANT: Droual, Anne-Marie
; APPLICANT: Mironov, Vladimir
; APPLICANT: Inz,, Dirk
; APPLICANT: Hatzfeld, Yves
; TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
; FILE REFERENCE: 1187-13
; CURRENT APPLICATION NUMBER: US/10/333,006
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: PCT/IB01/01492
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/218,471
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/241,219
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (32)..(32)
; OTHER INFORMATION: n = A, T, C or G
US-10-333-006-5
```

```
Alignment Scores:
Pred. No.: 3.12e-30 Length: 568
Score: 399.50 Matches: 95
Percent Similarity: 81.90% Conservative: 0
Best Local Similarity: 81.90% Mismatches: 14
Query Match: 30.64% Indels: 7
DB: 16 Gaps: 7
```

US-09-993-808B-2 (1-256) x US-10-333-006-5 (1-568)

```
QY 146 GlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGlySerGlyProAspArg 165
Db 8 GGGGATCAGTCTCCCGGAGCACTCGGGGAGCGTCCCGGACCGC 64
QY 166 GluArgArgGluThrThrPro---SerSerArgAlaHis---GlyGluLeuSerAspLeu 183
```


| | | | |
|----|-----|--|-----|
| Db | 65 | GAGAGGAGAGAGACGACGCCATCGTTCGAGCCGGCGCACGGCGGCGAGCTCAGCGATCTG | 124 |
| QY | 184 | GluserAspLeuAlaGly---HisLysThrGly---ProSerLeuProAla---AlaThr | 200 |
| Db | 125 | GAGTCGATCTGGTGGGCGGCAGAAAGACTGGCTGCTCGTCTGCGCGGCACAACAACA | 184 |
| QY | 201 | ProAlaAlaGluLeuileValProProAlaHisGluileGlnGluPheAlaAlaAala | 220 |
| Db | 185 | TCGGCTGCGAGCTGATCGTGCCGCCAGCACAGGAGATCCAGGAATCTTCGGGCGCC | 244 |
| QY | 221 | GluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyVal | 240 |
| Db | 245 | GAGGCGGCCCATGCCAAACGCTTTGCTTCCAAGTACAACTTCGACTTCGTCGCGGCGGTG | 304 |
| QY | 241 | ProLeuAspAlaGlyGlyArgPheGluTrpAlaProValValSerile | 256 |
| Db | 305 | CCGCTCGACGCC---GGCCGGTTCGAGTGGACGCCAGGGGTTCAGCATC | 349 |

RESULT 12
US-10-437-963-19949
; Sequence 19949, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 19949
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT MRT4530 2535C.1

| | |
|------------------------|----------|
| Alignment Scores: | |
| Pred. No.: | 1.02e-21 |
| Score: | 315.50 |
| Percent Similarity: | 46.62% |
| Best Local Similarity: | 37.37% |
| Query Match: | 24.19% |
| DB: | 17 |
| Length: | 1065 |
| Matches: | 105 |
| Conservative: | 26 |
| Mismatches: | 81 |
| Indels: | 69 |
| Gaps: | 13 |

Best Local Similarity: 36.39% Mismatches: 78
Query Match: 24.08% Indels: 82
DB: 17 Gaps: 14

US-09-993-808B-2 (1-256) x US-10-437-963-19960 (1-1398)

```
Qy 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAAGlyAlaGluValAlaAAlaValGlu 20
    |||||
Db 1272 ATGGGAAG---AAGAAGAAGCGGACGGCGCGGCGGAGGAGGAGGAGGCGGGTGGTG 1216
    |||||
Qy 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyVal 40
    |||||
Db 1215 GTCGGC-----GGCGTCCGTACGCGG----- 1195
    |||||
Qy 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaAlaValSer 60
    |||||
Db 1194 GCGCGCGTCACGGCGGAGGAGGTGTGTCGAGCGCGGAGGAGGAGGTGTGTTGGTGGGC 1135
    |||||
Qy 61 AlaGlyGlyAspGlyGlySer-----CysTyrIleHisLeu 72
    |||||
Db 1134 CGTGGCGGTGGCGGTGGCAGTGGCGGAGACGATGGCGAGGCGGATGCTATCTGCGTCTG 1075
    |||||
Qy 73 ArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAspSerValPro 92
    |||||
Db 1074 CGGAGCAGGAGGCTG----- 1060
    |||||
Qy 93 ThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGly 112
    |||||
Db 1059 ---CCCTTCGTGGCGCGCGGTGTGTCGTCGCGGAGGAGGAGGAGGCGCTCGGTGATTCTG 1003
    |||||
Qy 113 LeuSerArgCysSerSerThrAlaSerSer-----ValAsnLeuGlyLeuGlyGly 129
    |||||
Db 1002 GTGGCGGAGCGGCTTCGTGCTGTCGCGGCGGTGGAATTGTTGGGCTGTTCTGTT 943
    |||||
Qy 130 GlnArgGlySerHisThrCysArgSerTyrAspAlaAla----- 142
    |||||
Db 942 GAGGAGGAGGCTATGGCCGAGAGGTGATGATGAGCGCGCTTGGGATGTGTTGTTT 883
    |||||
Qy 143 -----GluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsn 158
    |||||
Db 882 CAGGTTTGACGACGAGGCGGAGGATCAC-----GACGAGGAGAGCTCCGTGCGCGAC 829
    |||||
Qy 159 SerGlySerGlyProAspArgGluArgGluThrThrProSerSerArg----- 175
    |||||
Db 828 TCCGGCTGCGGC-----CGCGAGAGAGCGCGACGACGCGCTGAGCCGCGCGCGCG 775
    |||||
Qy 176 AlaHisGlyGluLeuSerAspLeuGluSerAsp----- 186
    |||||
Db 774 GGAGACGGGACTCGAGCGACGCGGAGTCAACACGAGGAGGCCAAGCAGCAATGTGCGCG 715
    |||||
Qy 187 -----LeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThr 200
    |||||
Db 714 CGGAGTTCGACGACCTCAGCAGCTGCATTTACGCGGGA-----GCGACGACG 667
    |||||
Qy 201 ProAlaAlaGluLeuIleValProPro-----AlaHisGluIleGlnGluPhePheAla 218
    |||||
Db 666 AGGAGTTCAGGATGATGGCACCGCGCGCGCGGCGGCGGAGATCGAGGAGTTCTCTCGCC 607
    |||||
Qy 219 AlaAlaGluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArg 238
    |||||
Db 606 GCTGCGGAGAGGTCCGAGGCGCGAGCGCTTCGCGCGCAAGTACACTTCGACGTGTTGCGC 547
    |||||
Qy 239 GlyValProLeuAspAlaGly-----GlyArgPheGluTip 250
    |||||
Db 546 GCGGTGCGCTCGACGCGCGCGCGCGCGCGGCGGCGGTTTCAATGG 505
    |||||
```

RESULT 14

US-10-333-006-54
; Sequence 54, Application US/10333006
; Publication No. US20040019926A1
; GENERAL INFORMATION:
; APPLICANT: Frankard, Valerie Marie-No. US20040019926A1lle S.
; APPLICANT: Peres Bota, Adrian Marius

; APPLICANT: Droual, Anne-Marie
; APPLICANT: Mironov, Vladimir
; APPLICANT: Inz, Dirk
; APPLICANT: Hatzfeld, Yves
; TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
; FILE REFERENCE: 1187-13
; CURRENT APPLICATION NUMBER: US/10/333,006
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: PCT/IB01/01492
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/218,471
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/241,219
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54

; LENGTH: 681
; TYPE: DNA

; ORGANISM: Artificial sequence

; FEATURE:

; OTHER INFORMATION: probe or primer

; NAME/KEY: misc feature

; LOCATION: (355)..(356)

; OTHER INFORMATION: n = A, T, C or G

US-10-333-006-54

Alignment Scores:

Pred. No.: 2,38e-18 Length: 681
Score: 279.00 Matches: 98

Percent Similarity: 43.51% Conservative: 26

Best Local Similarity: 34.39% Mismatches: 71

Query Match: 21.40% Indels: 90

DB: 16 Gaps: 13

US-09-993-808B-2 (1-256) x US-10-333-006-54 (1-681)

```
Qy 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAAGlyAlaGluValAlaAAlaValGlu 20
    |||||
Db 1 ATGGGAAG---AAGAAGAAGCGGACGGCGCGGCGGAGGAGGAGGCGGGTGGTG 57
    |||||
Qy 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyVal 40
    |||||
Db 58 GTCGGC-----GGCGTCCGTACGCGG----- 78
    |||||
Qy 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaAlaValSer 60
    |||||
Db 79 GCGCGCGTCACGGCGGAGGAGGTGTGTCGCGAGCGCGGAGGAGGTGTGTTGGTGGGC 138
    |||||
Qy 61 AlaGlyGlyAspGlyGlySer-----CysTyrIleHisLeu 72
    |||||
Db 139 CGTGGCGGTGGCGGTGGCAGTGGCGGAGACGATGGCGAGGCGGATGCTATCTGCGTCTG 198
    |||||
Qy 73 ArgSerArgMetLeuPheMetAlaProProGlnProGlnProSerValAspSerValPro 92
    |||||
Db 199 CGGAGCAGGAGGCTG----- 213
    |||||
Qy 93 ThrProValGluAlaAlaAspGlyAlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGly 112
    |||||
Db 214 ---CCCTTCGTGGCGCGCGGTGTGTCGTCGCGGAGGAGGAGGCGCTCGGTGATTCTG 270
    |||||
Qy 113 LeuSerArgCysSerSerThrAlaSerSerValAsnLeuGlyLeuGlyGlnArgGly 132
    |||||
Db 271 GTGGCGGAGGCGGCTTCGTGCTGTCGTCGCGCGGCGGTGGAATTGTTGGGC----- 321
    |||||
Qy 133 SerHisThrCysArgSerTyrAspAlaAlaGluAlaGlyGlyAspHisValLeuValAsp 152
    |||||
Db 322 -----TGT-----TCTGTCGAGGAGGAGGAGGATGTCGCGGAG 351
    |||||
Qy 153 ValSerAlaAlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThrPro 172
    |||||
Db 352 AAGNNGAGCGCG-----ACGACGCGCG 372
    |||||
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2004, 12:44:58 ; Search time 3090 Seconds
(without alignments)
3590.879 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 256
Sequence: 1 MGKYMRCRGAAGAEVAVE.....VRGVPLDAGGRFEWAPVWSI 256

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues
Word size: 1

Total number of hits satisfying chosen parameters: 6934743

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US0993808/runat_01102004_171726_13707/app_query.fasta_1.455
-DB=GenEmbl -QFMT=fastap -SUFFIX=oligo.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0993808@cgn_1_1_5600@runat_01102004_171726_13707 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 21 | 8.2 | 568 | 6 | AX406682 Sequence |
| 2 | 17 | 6.6 | 1242 | 6 | AX406686 Sequence |
| 3 | 17 | 6.6 | 1266 | 8 | AK103084 Oryza sat |
| C 4 | 17 | 6.6 | 135670 | 8 | AP005002 Oryza sat |
| 5 | 11 | 4.3 | 187157 | 9 | AL139384 Human DNA |
| 6 | 11 | 4.3 | 212007 | 2 | AC111297 Rattus no |
| 7 | 11 | 4.3 | 224600 | 2 | AC111313 Rattus no |
| 8 | 10 | 3.9 | 681 | 6 | AX406731 Sequence |
| 9 | 10 | 3.9 | 1073 | 6 | AX406733 Sequence |
| C 10 | 10 | 3.9 | 2347 | 6 | AX835276 Sequence |
| C 11 | 10 | 3.9 | 2347 | 9 | AK098250 Homo sapi |
| 12 | 10 | 3.9 | 3525 | 6 | AX654623 Sequence |
| C 13 | 10 | 3.9 | 10562 | 1 | AE004870 Pseudomon |
| C 14 | 10 | 3.9 | 12639 | 1 | AE005072 Halobacte |
| C 15 | 10 | 3.9 | 110000 | 8 | Continuation (23 o |
| C 16 | 10 | 3.9 | 151208 | 2 | AP003525 Oryza sat |
| C 17 | 10 | 3.9 | 156654 | 8 | AC074283 Oryza sat |
| C 18 | 10 | 3.9 | 166660 | 9 | AC008073 Homo sapi |
| C 19 | 10 | 3.9 | 187916 | 2 | AP003458 Oryza sat |
| C 20 | 10 | 3.9 | 210614 | 1 | AB088224 Streptomy |
| C 21 | 10 | 3.9 | 247105 | 10 | AC111653 Rattus no |
| C 22 | 10 | 3.9 | 302085 | 8 | AE017054 Oryza sat |
| C 23 | 9 | 3.5 | 228 | 9 | AY270716 Homo sapi |
| C 24 | 9 | 3.5 | 268 | 9 | AY270729 Homo sapi |
| 25 | 9 | 3.5 | 338 | 9 | HSA1322B9 |
| 26 | 9 | 3.5 | 350 | 8 | HVJ222778 |
| 27 | 9 | 3.5 | 425 | 9 | AY270710 Homo sapi |
| 28 | 9 | 3.5 | 425 | 9 | AY270722 Homo sapi |
| 29 | 9 | 3.5 | 436 | 6 | AR262501 Sequence |
| 30 | 9 | 3.5 | 436 | 6 | AX022310 Sequence |
| 31 | 9 | 3.5 | 436 | 6 | AX030777 Sequence |
| 32 | 9 | 3.5 | 436 | 6 | BD007462 Remedies |
| 33 | 9 | 3.5 | 493 | 6 | AX406678 Sequence |
| 34 | 9 | 3.5 | 688 | 8 | AF205647 Gomphidiu |
| 35 | 9 | 3.5 | 771 | 6 | AR387547 Sequence |
| 36 | 9 | 3.5 | 792 | 6 | AR390079 Sequence |
| 37 | 9 | 3.5 | 975 | 6 | AX573934 Sequence |
| 38 | 9 | 3.5 | 1601 | 1 | AB021882 Streptomy |
| 39 | 9 | 3.5 | 1716 | 1 | AB009289 Streptomy |
| 40 | 9 | 3.5 | 1776 | 1 | AF455754 Bordetell |
| C 41 | 9 | 3.5 | 2043 | 10 | AF121907 Mus muscu |
| C 42 | 9 | 3.5 | 2113 | 6 | AX834764 Sequence |
| C 43 | 9 | 3.5 | 2113 | 9 | AK097418 Homo sapi |
| C 44 | 9 | 3.5 | 2163 | 10 | BC017597 Mus muscu |
| 45 | 9 | 3.5 | 2279 | 8 | AK100849 Oryza sat |

ALIGNMENTS

AX406682
LOCUS AX406682 568 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 5 from Patent WO0228893.
ACCESSION AX406682
VERSION AX406682.1 GI:21439630
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1
REFERENCE Droual,A.-M., Frankard,V.M.-N., Inze,D., Mironov,V., Peres,B.A.D.
AUTHORS and Hatzfeld,Y.
TITLE Plant cyclin-dependent kinase inhibitors
JOURNAL Patent: WO 0228893-A 5 11-APR-2002;
CROPDESIGN N V (BE)
FEATURES Location/Qualifiers
source 1..568
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Alignment Scores:
Pred. No.: 1.02e-07 Length: 568
Score: 21.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.20% Indels: 0
DB: 6 Gaps: 0

US-09-993-808B-2 (1-256) x AX406682 (1-568)

Qy 225 AlAlysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyValProLeuAspAla 244
|||||
Db 257 GCCAAGCGCTTGCTTCCAAAGTACAACTTCGACTTCGCGCGGTGCCCTCGACGCC 316
Qy 245 Gly 245
|||
Db 317 GGC 319

RESULT 2
AX406686
LOCUS AX406686 1242 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 9 from Patent WO0228893.
ACCESSION AX406686
VERSION AX406686.1 GI:21439634
KEYWORDS
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1
REFERENCE Droual,A.-M., Frankard,V.M.-N., Inze,D., Mironov,V., Peres,B.A.D.
AUTHORS and Hatzfeld,Y.
TITLE Plant cyclin-dependent kinase inhibitors
JOURNAL Patent: WO 0228893-A 9 11-APR-2002;
CROPDESIGN N V (BE)

FEATURES Location/Qualifiers
source 1..1242
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"

ORIGIN

Alignment Scores:
Pred. No.: 0.000289 Length: 1242
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.64% Indels: 0

DB: 6 Gaps: 0
US-09-993-808B-2 (1-256) x AX406686 (1-1242)
Qy 20 GluValThrGlnValGlyValArgThrArgSerArgSerAlaAla 36
|||||
Db 75 GAGGTCACGACGGTGGTGGCGTCGGACGAGGTCGAGGTCGCGCGCG 125
RESULT 3
AK103084
LOCUS AK103084 1266 bp mRNA linear PLN 24-JUL-2003
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J033118G18, full
insert sequence.
ACCESSION AK103084
VERSION AK103084.1 GI:32988293
KEYWORDS FLI_CDNA; CAP trapper.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE
AUTHORS

1 The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-Length cDNA Project Team:
Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group: Ohtsuki,Y., Murakami,K.,
Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,
Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,
Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN:
Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,
Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,
Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,
Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,
Yoshino,M. and Hayashizaki,Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice

TITLE

JOURNAL Science 301 (5631), 376-379 (2003)

MEDLINE 22752273

PUBMED 12869764

2 (bases 1 to 1266)

REFERENCE

AUTHORS

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,
Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,
Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imamura,K.,
Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,
Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M.,
Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,
Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M.,
Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,
Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M.,
Namiki,T., Narikawa,R., Niikura,J., Nishi,K., Nomura,K.,
Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,
Osato,N., Ota,Y., Ohtsuki,K., Ohtsuki,K., Oka,M., Ooka,H.,
Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K.,
Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S.,
Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y.,
Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A.,
Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,W.,
Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and
Yoshimura,A.

Direct Submission

TITLE

JOURNAL

Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT

This clone is one of the 28K full-length cDNA clones from japonica

rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and
Yamamoto, M.
FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,
Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.
Location/Qualifiers
1. 1266
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J033118G18"

ORIGIN

Alignment Scores:
Pred. No.: 0.000293 Length: 1266
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.64% Indels: 0
DB: 8 Gaps: 0

US-09-993-808B-2 (1-256) x AK103084 (1-1266)

Qy 20 GlnValThrGlnValGlyValArgThrArgSerArgSerAlaAla 36
|||||
Db 118 GAGTACGCAGGTGGTTGGCTCGGACGAGGTGCGGCGCGCGGCG 168

RESULT 4
AP005002/c
LOCUS AP005002 135670 bp DNA linear PLN 14-AUG-2003
DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
PAC clone:P0486G03, complete sequence.
ACCESSION AP005002
VERSION AP005002.2 GI:33636434
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1
REFERENCE
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC
clone:P0486G03
JOURNAL Published Only in Database (2002)
REFERENCE 2 (bases 1 to 135670)
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Direct Submission
JOURNAL Submitted (27-MAR-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai

2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Aug 13, 2003 this sequence version replaced gi:19773538.
The orientation of the sequence is from T7 to SP6 of the PAC clone.
Location/Qualifiers
1. 135670
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="P0486G03"

ORIGIN

Alignment Scores:
Pred. No.: 0.00901 Length: 135670
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.64% Indels: 0
DB: 8 Gaps: 0

US-09-993-808B-2 (1-256) x AP005002 (1-135670)

Qy 20 GlnValThrGlnValGlyValArgThrArgSerArgSerAlaAla 36
|||||
Db 80138 GAGTACGCAGGTGGTTGGCTCGGACGAGGTGCGGCGCGGCG 80088

RESULT 5
AL139384
LOCUS AL139384 187157 bp DNA linear PRI 20-JUN-2001
DEFINITION Human DNA sequence from clone RP11-88E10 on chromosome 13q33.1-34,
complete sequence.
ACCESSION AL139384
VERSION AL139384.17 GI:14530206
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 187157)
AUTHORS Smith, M.
TITLE Direct Submission
JOURNAL Submitted (20-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
requests: clonerequest@sanger.ac.uk
On Jun 22, 2001 this sequence version replaced gi:11610962.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 13, constructed by the Sanger Centre Chromosome 13
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr13
RP11-88E10 is from the library RPCI-11.1 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. IMPORTANT: This sequence is not the entire insert of clone RP11-88E10. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.
The true left end of clone RP11-88E10 is at 1 in this sequence. The true left end of clone RP11-480K16 is at 187058 in this sequence. The true right end of clone RP11-144L2 is at 50570 in this sequence.

FEATURES

source

Location/Qualifiers

1. .187157
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="13"
/map="q33.1-34"
/clone="RP11-88E10"
/clone_lib="RPCT-11.1"
1046. .1077
/note="MER75 repeat: matches 333. .364 of consensus"
1494. .1517
/note="12 copies 2 mer ac 100% conserved"
1520. .1818
/note="AluSg repeat: matches 1. .298 of consensus"
2127. .3461
/note="CpG island"
/evidence=not_experimental
3035. .3126
/note="46 copies 2 mer cc 64% conserved"
3136. .3177
/note="6 copies 7 mer gggcaga 83% conserved"
complement(4293. .4718)
/note="match: STS: Em:G29775"
4299. .4550
/note="LTR33 repeat: matches 256. .516 of consensus"
5281. .5623
/note="MLT1A1 repeat: matches 1. .365 of consensus"
5892. .6094
/note="MLT1H repeat: matches 79. .291 of consensus"
7071. .7380
/note="AluSx repeat: matches 1. .310 of consensus"
7620. .7668
/note="MER5B repeat: matches 1. .49 of consensus"
7692. .7953
/note="AluJo repeat: matches 1. .285 of consensus"
complement(7708. .8008)
/note="match: GSS: Em:AQ757664"
7993. .8051
/note="MER5B repeat: matches 40. .104 of consensus"
8068. .8448
/note="match: GSS: Em:AQ014714"
complement(8224. .8765)
/note="match: GSS: Em:AQ141652"
complement(8600. .8786)
/note="match: STS: Em:G50946
match: GSS: Em:B82990"
complement(8612. .8786)
/note="match: GSS: Em:B83029"
9048. .9133
/note="2 copies 43 mer 94% conserved"
9922. .10225
/note="AluSg repeat: matches 1. .305 of consensus"
10390. .10705
/note="L1MB3 repeat: matches 5870. .6182 of consensus"
10731. .11023
/note="AluY repeat: matches 1. .293 of consensus"
11112. .11357
/note="AluSg/x repeat: matches 70. .301 of consensus"
11359. .11644
/note="AluJb repeat: matches 1. .286 of consensus"
13908. .14343
/note="SVA repeat: matches 956. .1385 of consensus"
14344. .15086
/note="SVA repeat: matches 6. .760 of consensus"

repeat_region
15147. .15443
/note="AluSx repeat: matches 1. .295 of consensus"
15895. .16203
/note="AluSq repeat: matches 1. .310 of consensus"
16519. .16878
/note="HAL1 repeat: matches 721. .1114 of consensus"
complement(18798. .19280)
/note="match: GSS: Em:AQ147060"
19105. .19481
/note="MLT1E repeat: matches 176. .568 of consensus"
19485. .19784
/note="AluSx repeat: matches 2. .302 of consensus"
complement(19635. .20146)
/note="match: GSS: Em:AQ395327"
19822. .20067
/note="6 copies 41 mer 65% conserved"
19855. .20082
/note="12 copies 19 mer 64% conserved"
19857. .20072
/note="8 copies 27 mer 63% conserved"
19911. .20082
/note="4 copies 43 mer 70% conserved"
20932. .21232
/note="AluSx repeat: matches 1. .302 of consensus"
complement(21053. .21546)
/note="match: GSS: Em:AQ512552"
complement(21119. .21534)
/note="match: GSS: Em:AQ564752"
21393. .21568
/note="MIR repeat: matches 36. .223 of consensus"
complement(21587. .21880)
/note="match: STS: Em:HSPH07E6"
21789. .21843
/note="TAR1 repeat: matches 635. .686 of consensus"
22130. .22437
/note="AluSx repeat: matches 1. .308 of consensus"
23147. .23486
/note="match: GSS: Em:AQ080832"
23229. .23508
/note="AluSx repeat: matches 1. .304 of consensus"
complement(23618. .24039)
/note="match: GSS: Em:AQ470695"
24469. .24761
/note="AluSx repeat: matches 1. .294 of consensus"
24884. .25757
/note="L1M4 repeat: matches 3366. .4222 of consensus"
25960. .27013
/note="527 copies 2 mer ac 69% conserved"
27026. .27191
/note="83 copies 2 mer ca 72% conserved"
28259. .28394
/note="MERSA repeat: matches 29. .189 of consensus"
28938. .29244
/note="AluSq repeat: matches 1. .298 of consensus"
29249. .29548
/note="AluSx repeat: matches 1. .302 of consensus"
29676. .30494
/note="L1MA6 repeat: matches 4193. .5002 of consensus"
30495. .30795
/note="AluY repeat: matches 2. .302 of consensus"
30796. .31938
/note="L1MA6 repeat: matches 5002. .6255 of consensus"
31941. .32000
/note="MER41B repeat: matches 483. .547 of consensus"
32003. .32059
/note="L1MA5 repeat: matches 6242. .6300 of consensus"
32103. .32448
/note="TIGGER2 repeat: matches 1. .357 of consensus"
32449. .32758
/note="AluSx repeat: matches 1. .311 of consensus"
32759. .32822
/note="TIGGER2 repeat: matches 357. .422 of consensus"
32823. .33132

```
/note="AluY repeat: matches 2. .310 of consensus"
33133. .33489
/note="TIGGER2 repeat: matches 422. .744 of consensus"
33553. .33764
/note="TIGGER2 repeat: matches 2462. .2690 of consensus"
34943. .35023
/note="MER81 repeat: matches 7. .88 of consensus"
complement(35414. .35877)
/note="match: GSS: Em:AQ220607"
complement(35464. .35877)
/note="match: GSS: Em:AQ451480"
35690. .35727

Alignment Scores:
Pred. No.: 726 Length: 187157
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.30% Indels: 0
DB: 9 Gaps: 0

US-09-993-808B-2 (1-256) x AL139384 (1-187157)
QY 107 AlaAlaLeuAlaAaAGlyLeuSerArgCysSer 117
|||||
Db 68751 GCAGCGCTGGCGCAGGACTGAGCGGTGCTCC 68783

RESULT 6
AC111297 212007 bp DNA linear HTG 13-MAY-2003
LOCUS Rattus norvegicus clone CH230-223J18, *** SEQUENCING IN PROGRESS
DEFINITION *** 3 unordered pieces.
ACCESSION AC111297
VERSION AC111297.5 GI:30578726
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 212007)
Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Bliswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwackeleme,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
```

```
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 212007)
Worley,K.C.
Direct Submission
Submitted (19-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 212007)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:25007665.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GLYZ
Center clone name: CH230-223J18
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 192461 bases at least Q40
Consensus quality: 196409 bases at least Q30
Consensus quality: 199024 bases at least Q20
Estimated insert size: 203643; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 208729: contig of 208729 bp in length
* 208730 208829: gap of unknown length
* 208830 210590: contig of 1761 bp in length
```

* 210591 210690: gap of unknown length
* 210691 212007: contig of 1317 bp in length.
FEATURES
source Location/Qualifiers
1. 212007
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-223J18"
143929. .146581
/note="wgs_contig"
154185. .156685
/note="wgs_contig"

misc_feature

/note="wgs_contig"

154185. .156685

/note="wgs_contig"

ORIGIN

Alignment Scores:
Pred. No.: 795 Length: 212007
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.30% Indels: 0
DB: 2 Gaps: 0

US-09-993-808B-2 (1-256) x AC1111297 (1-212007)

QY 107 AlaAlaLeuAlaAAGlyLeuSerArgCysSer 117

|||||

Db 134431 GCTGCTTGGCTGCAGGCTGAGTGGTCTCT 134463

RESULT 7

AC111313

LOCUS AC111313 224600 bp DNA linear HTG 13-MAY-2003
DEFINITION Rattus norvegicus clone CH230-243L23, WORKING DRAFT SEQUENCE, 3
unordered pieces.

ACCESSION

AC111313.9 GI:30578774

VERSION HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

KEYWORDS

Rattus norvegicus (Norway rat)

SOURCE

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 224600)

REFERENCE
Mutzker, M., Lee, A., Adams, C., Alder, J.,
Allen, D., Marie, D., Metzger, M., Brooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Duxbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flag, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebegeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensuhewa, L., Loulseg, H., Lozada, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,

Nwaakelemeh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkoch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L.,
Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steimle, M., Strong, N., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.

TITLE

JOURNAL

Direct Submission
Unpublished
2 (bases 1 to 224600)

AUTHORS

TITLE

JOURNAL

Submitted (19-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 224600)

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:23107935.

The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

COMMENT

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GLZY

Center clone name: CH230-243L23

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 209953 bases at least Q40

Consensus quality: 210980 bases at least Q30

Consensus quality: 211844 bases at least Q20

Estimated insert size: 226554; sum-of-contigs estimation

Quality coverage: 9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 3 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 221276: contig of 221276 bp in length

* 221277 221376: gap of unknown length

* 221377 222686: contig of 1310 bp in length
* 222687 222786: gap of unknown length
* 222787 224600: contig of 1814 bp in length.

FEATURES
source
1. .224600
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-243L23"

ORIGIN

Alignment Scores:
Pred. No.: 829 Length: 224600
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.30% Indels: 0
DB: 2 Gaps: 0

US-09-993-808B-2 (1-256) x AC1111313 (1-224600)

QY 107 AlaAlaLeuAlaAlaGlyLeuSerArgCysSer 117
|||||

Db 143525 GCTGCTTGGCTGCAGGGCTGAGTCGCTGCTCT 143557
|||||

RESULT 8

AX406731 LOCUS AX406731 681 bp DNA linear PAT 14-JUN-2002

DEFINITION Sequence 54 from Patent WO0228893.

ACCESSION AX406731

VERSION AX406731.1 GI:21439657

KEYWORDS

SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.

REFERENCE

AUTHORS

Droual,A.-M., Frankard,V.M.-N., Inze,D., Mironov,V., Peres,B.A.D.
and Hatzfeld,Y.

TITLE Plant cyclin-dependent kinase inhibitors

JOURNAL Patent: WO 0228893-A 54 11-APR-2002;

CROPDESIGN N V (BE)

FEATURES

source
1. .681
/organism="synthetic construct"
/mol_type="synthetic construct"
/db_xref="taxon:32630"
/note="probe or primer"

ORIGIN

Alignment Scores:
Pred. No.: 74.6 Length: 681
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 6 Gaps: 0

US-09-993-808B-2 (1-256) x AX406731 (1-681)

QY 237 ValArgGlyValProLeuAspAlaGlyGly 246
|||||

Db 613 GTGCGCGGCTGCCGCTCGACGCGGCGGC 642
|||||

RESULT 9

AX406733

LOCUS AX406733 1073 bp DNA linear PAT 14-JUN-2002

DEFINITION Sequence 56 from Patent WO0228893.

ACCESSION AX406733

VERSION AX406733.1 GI:21439658

KEYWORDS

SOURCE Oryza sativa

ORGANISM

Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE

AUTHORS

Droual,A.-M., Frankard,V.M.-N., Inze,D., Mironov,V., Peres,B.A.D.
and Hatzfeld,Y.

TITLE Plant cyclin-dependent kinase inhibitors

JOURNAL Patent: WO 0228893-A 56 11-APR-2002;

CROPDESIGN N V (BE)

FEATURES

source
1. .1073
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"

ORIGIN

Alignment Scores:
Pred. No.: 104 Length: 1073
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 6 Gaps: 0

US-09-993-808B-2 (1-256) x AX406733 (1-1073)

QY 237 ValArgGlyValProLeuAspAlaGlyGly 246
|||||

Db 1005 GTGCGGGGTGCCGCTCGACGCGGCGGC 1034
|||||

RESULT 10

AX835276/c

LOCUS AX835276 2347 bp DNA linear PAT 15-DEC-2003

DEFINITION Sequence 2400 from Patent EP1347046.

ACCESSION AX835276

VERSION AX835276.1 GI:39921411

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and
Masuho,Y.

TITLE Full-length cDNA sequences

JOURNAL Patent: EP 1347046-A 2400 24-SEP-2003;

Research Association for Biotechnology (JP)

FEATURES

source
1. .2347
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 185 Length: 2347
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 6 Gaps: 0

US-09-993-808B-2 (1-256) x AX835276 (1-2347)

QY 24 ValValGlyValArgThrArgSerArgSer 33
|||||

Db 66 GTCGTAGGCGTCCGCACCGCTCTCGCTCA 37
|||||

RESULT 11

AK098250/c

LOCUS

DEFINITION

Homo sapiens cDNA FLJ40931 fis, clone UTERU2007004.
AK098250 2347 bp mRNA linear PRI 15-JUL-2002

ACCESSION AK098250
VERSION AK098250.1 GI:21758226
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K.,
Arita, M., Musashino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R.,
Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
Wagatsuma, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B.,
Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and
Isogai, T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2347)
AUTHORS Isogai, T. and Yamamoto, J.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
source Location/Qualifiers
1. 2347
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UTERU2007004"
/tissue type="uterus"
/clone_lib="UTERU2"
/note="cloning vector: pME18SFL3"
ORIGIN
Alignment Scores: 185 Length: 2347
Pred. No.: 10.00 Matches: 10
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 3.91% Gaps: 0
DB: 9
US-09-993-808B-2 (1-256) x AK098250 (1-2347)
QY 24 ValValGlyValArgThrArgSerArgSer 33
|||||
Db 66 GTCTAGCGCTCCGCCACCCGCTCTCGCTCA 37
RESULT 12
AX654623
LOCUS AX654623 3525 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 4493 from Patent WO03000898.
ACCESSION AX654623
VERSION AX654623.1 GI:29157437
KEYWORDS
SOURCE
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1
REFERENCE
AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
TITLE Plant genes involved in defense against pathogens

JOURNAL Patent: WO 03000898-A 4493 03-JAN-2003;
Syngenta Participations AG (CH)
FEATURES
source Location/Qualifiers
1. 3525
/organism="Oryza sativa"
/mol_type="unassigned DNA"
/db_xref="taxon:4530"
ORIGIN
Alignment Scores: 249 Length: 3525
Pred. No.: 10.00 Matches: 10
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 3.91% Gaps: 0
DB: 6
US-09-993-808B-2 (1-256) x AX654623 (1-3525)
QY 28 ArgThrArgSerArgSerAlaAlaAlaThr 37
|||||
Db 206 CGTAGCGCTCGCGCTCTGCGCGGCGACG 235
RESULT 13
AE004870/c
LOCUS AE004870 10562 bp DNA linear BCT 19-FEB-2003
DEFINITION Pseudomonas aeruginosa PAO1, section 431 of the complete
genome.
ACCESSION AE004870 AE004091
VERSION AE004870.1 GI:9950807
KEYWORDS
SOURCE Pseudomonas aeruginosa PAO1
ORGANISM Pseudomonas aeruginosa PAO1
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 10562)
AUTHORS Stover, C.K., Pham, X.-Q.T., Erwin, A.L., Mizoguchi, S.D., Warren, P.,
Hickey, M.J., Brinkman, F.S.L., Hufnagle, W.O., Kowalik, D.J.,
Lagrou, M., Garber, R.L., Goltry, L., Tolentino, E.,
Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N.,
Folger, K.R., Kas, A., Larbig, K., Lim, R.M., Smith, K.A., Spencer, D.H.,
Wong, G.K.-S., Wu, Z. and Paulsen, I.T.
TITLE Complete genome sequence of Pseudomonas aeruginosa PAO1, an
opportunistic pathogen
JOURNAL Nature 406 (6799), 959-964 (2000)
MEDLINE 20437337
PUBMED 10984043
REFERENCE 2 (bases 1 to 10562)
AUTHORS Stover, C.K., Pham, X.-Q.T., Erwin, A.L., Mizoguchi, S.D., Warren, P.,
Hickey, M.J., Brinkman, F.S.L., Hufnagle, W.O., Kowalik, D.J.,
Lagrou, M., Garber, R.L., Goltry, L., Tolentino, E.,
Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N.,
Folger, K.R., Kas, A., Larbig, K., Lim, R.M., Smith, K.A., Spencer, D.H.,
Wong, G.K.-S., Wu, Z., Paulsen, I.T., Reizer, J., Saier, M.H.,
Hancock, R.E.W., Lory, S. and Olson, M.V.
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2000) Department of Medicine and Genetics,
University of Washington Genome Center, University Of Washington,
Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 10562)
Pseudomonas aeruginosa Community Annotation Project (PseudoCAP)
Direct Submission
Submitted (04-FEB-2003) Department of Molecular Biology and
Biochemistry, Simon Fraser University, 8888 University Dr.,
Burnaby, British Columbia V5A 1S6, Canada

This represents the February 3, 2003 version of the continually
updated, reviewed, Pseudomonas aeruginosa PAO1 genome annotation,
from PseudoCAP (see <http://www.pseudomonas.com> for latest updates
and links to alternate annotations). PseudoCAP is coordinated by
Fiona S.L. Brinkman (Simon Fraser University, Canada) and Robert
E.W. Hancock (University of British Columbia, Canada). We welcome

submission through www.pseudomonas.com of any proposed changes.

'Protein name confidence' is used to rate our confidence of the accuracy of the protein name.
Class 1: Function experimentally demonstrated in *P. aeruginosa*.
Class 2: Function of highly similar gene experimentally demonstrated in another organism (and gene context consistent in terms of pathways its involved in, if known).
Class 3: Function proposed based on presence of conserved amino acid motif, structural feature or limited sequence similarity to an experimentally studied gene.
Class 4: Homologs of previously reported genes of unknown function, or no similarity to any previously reported sequences.

FEATURES

```
source
1. .10562
/organism="Pseudomonas aeruginosa PAO1"
/mol_type="genomic DNA"
/strain="PAO1"
/db_xref="taxon:208964"
complement(79. .1617)
/locus_tag="PA4562"
/note="synonym: mvnN"
complement(79. .1617)
/locus_tag="PA4562"
/note="Protein name confidence: Class 4 (homologs of previously reported genes of unknown function, or no similarity to any previously reported sequences)"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAG07950.1"
/db_xref="GI:9950808"
/translation="MNLLKSLAAVSSITMLSRVLGFVRDTILARIFGAGLATDAFFVA
FKPLNLLRRIFAFAGAFSAFVPILAEYKNOQGEATRTFIAYVSGLLTLVLALVTALG
ILAAPVWIVTAPGADTPKEFALTDDLRTVPYILLISLSSLAGAILTNWRFSPV
AFVPTLLNVAMIGFALETPYFDPVPMVLGAWLAGLLQLLYQLPHERKIGMLVLP
LNLRDSGVWRVWMLMPLAILGVSVQISLIINTIFASFLAAGSVWYVADRIMELPS
GVLGALGTILPLMAKTYSNKDRHEYSRLDGLRLCFLVLVPCSLAAILAEPLTV
SLFOYKFTTVDAANTQRALVAYSVGLLGIILVKVLAPGFYAQQNIRTPVKIALFTLV
STQLMNAFIPQLHAGLALSIGLAACLNAGLLYWQLRKQRLYLPQPGWAKFLVKLVV
AVLAMSAVLLATMHWLPAAWEQGMALERFLRLGLLVVAGLLAYFGMLALLGFRLRDFS
RAVL"
1901. .2176
/gene="rpst"
/locus_tag="PA4563"
1901. .2176
/gene="rpst"
/locus_tag="PA4563"
/note="Protein name confidence: Class 2 (high similarity to functionally studied protein)"
/codon_start=1
/transl_table=11
/product="30S ribosomal protein S20"
/protein_id="AAG07951.1"
/db_xref="GI:9950809"
/translation="WANTPSAKKRAKQAEKRRSHNASLRSMVRTYIKNVVKAIDAKDL
EKAQAFTAAVPVIDRMADKGIHKNKAARHKSRLSGHICALSTAAA"
complement(2243. .2707)
/locus_tag="PA4564"
/note="synonym: creA"
complement(2243. .2707)
/locus_tag="PA4564"
/note="Protein name confidence: Class 4 (homologs of previously reported genes of unknown function, or no similarity to any previously reported sequences)"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAG07952.1"
/db_xref="GI:9950810"
/translation="MRKLAGLLGGLLLPALAGABEIGEIVSTVFKWVGPNDKIVVEAF
DDPKVQGVTCYLSRAKTGVKGLGLAEDRAEAESIACRQVGPIQFSGELKDGEVFKQ
```

gene

CDS

```
RTSLVFKSMQVVRFFDRKRNALVYLVSVDRIEGSPQNAVTAIPLMPWPTAR"
complement(2719. .3837)
/gene="proB"
/locus_tag="PA4565"
complement(2719. .3837)
/gene="proB"
/locus_tag="PA4565"
/note="Protein name confidence: Class 2 (high similarity to functionally studied protein); Subcellular localization: Cytoplasmic; Subcellular localization confidence: Class 2"
/codon_start=1
/transl_table=11
/product="glutamate 5-kinase"
/protein_id="AAG07953.1"
/db_xref="GI:9950811"
/translation="MRDKVTGARRWVVKIGSALLTADGRGLDRNAMVWVEQMVALHC
AGIELVLVSSGAAGAAGMSRLGWVSRPSPMSHMLQAAASVQMGVLQAWESFALHGLQT
AQVLLTHDDLSDRKRYLNARSTLRTLVELGVVFINENDTVVTDEIRFGDNDTLAALV
ANLVEADLLVILTRDGMFDADPRNPDQALIYEARADDPQLDAVAGSAGALGRGGM
QTKLRAARLAARSQGHVTIVGGRIERVLDRLRAGERLGTLLTPDRSRKAARKQWLGH
LQMRGTLVLDGGAOKAVSQDHKSLLPVGKAVQGSFRRGEMVVCVDQGGREVARGLVN
YSALEAQKILGQPTDAIEALLGYVDGPVELVHRDNLVLV"
complement(3908. .5128)
/gene="obg"
/locus_tag="PA4566"
complement(3908. .5128)
/gene="obg"
/locus_tag="PA4566"
/note="Protein name confidence: Class 2 (high similarity to functionally studied protein)"
/codon_start=1
/transl_table=11
/product="GTP-binding protein Obg"
/protein_id="AAG07954.1"
/db_xref="GI:9950812"
/translation="MKFVDEVSIHVKAGDGGNGIMSFRRKFIKGGPNGGDGGGS
IYLEADVNLTLVDYRYTRRFDAQRGENGSKDCTGAKGDDLILPVPVGTVIDANTQ
EIIGDLTEPGQRLMVAQGGHGLGNTFRKFSSTNAPRQTTGPKGEARDLKLKVL
DVGLLGLPNAGKSTFIRAVSAKPKVADYPTTLVPLNGLSVGVRYKSFVVADIPLLI
EGAAEGAGLGRFLKHLARTLRLHLVDMAPLDES DPADAAEVIIVRELGRFSALTER
ERWLVLNKMDOIIPAEAREARKQAVIERLWEGVPVVISALERDGTALSQDIMRYLD
ERTLRLEEDPQYAEELAEELDRRIEDEARARLQALDDARALRRSGLKNAGAVDDDDFDD
EEDDGDGPEIFYVP"
complement(5270. .5527)
/gene="rpmA"
/locus_tag="PA4567"
complement(5270. .5527)
/gene="rpmA"
/locus_tag="PA4567"
/note="Protein name confidence: Class 2 (high similarity to functionally studied protein)"
/codon_start=1
/transl_table=11
/product="50S ribosomal protein L27"
/protein_id="AAG07955.1"
/db_xref="GI:9950813"
/translation="MAHKKAGGSTRNGRDSKRLGKLVFGGQAVKAGNILVRQRTK
FHAGYGVGLGKDHTLFAKVDGVVKFETKGAFGKYSIVAA"
complement(5551. .5862)
/gene="rplU"
/locus_tag="PA4568"
complement(5551. .5862)
/gene="rplU"
/locus_tag="PA4568"
/note="Protein name confidence: Class 2 (high similarity to functionally studied protein)"
/codon_start=1
/transl_table=11
/product="50S ribosomal protein L21"
/protein_id="AAG07956.1"
/db_xref="GI:9950814"
/translation="MYAVITVGGKQHKVTEGTEFLKVEKLDVATGEAIDFDRVLLVANG
```

gene

CDS

```
complement(5270. .5527)
/gene="rpmA"
/locus_tag="PA4567"
complement(5270. .5527)
/gene="rpmA"
/locus_tag="PA4567"
/note="Protein name confidence: Class 2 (high similarity to functionally studied protein)"
/codon_start=1
/transl_table=11
/product="50S ribosomal protein L27"
/protein_id="AAG07955.1"
/db_xref="GI:9950813"
/translation="MAHKKAGGSTRNGRDSKRLGKLVFGGQAVKAGNILVRQRTK
FHAGYGVGLGKDHTLFAKVDGVVKFETKGAFGKYSIVAA"
complement(5551. .5862)
/gene="rplU"
/locus_tag="PA4568"
complement(5551. .5862)
/gene="rplU"
/locus_tag="PA4568"
/note="Protein name confidence: Class 2 (high similarity to functionally studied protein)"
/codon_start=1
/transl_table=11
/product="50S ribosomal protein L21"
/protein_id="AAG07956.1"
/db_xref="GI:9950814"
/translation="MYAVITVGGKQHKVTEGTEFLKVEKLDVATGEAIDFDRVLLVANG
```

EDVKIGLPVVEGAKVTAEVVSHGRHDKVRLIKFRRRKHHMKRQCHRWTFEIKITGIQ
A"
6103. .7071
/gene="ispB"
/locus_tag="PA4569"
/note="synonym: cel"

gene

Alignment Scores:
Pred. No.: 557 Length: 10562
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 1 Gaps: 0

US-09-993-808B-2 (1-256) x AE004870 (1-10562)

QY 113 LeuSerArgCysSerSerThrAlaSerSer 122
|||||
DB 620 CTGTCTCGCTGTTCCAGTACGGCAAGTTCA 591

RESULT 14

AE005072/c
LOCUS AE005072 12639 bp DNA linear BCT 12-FEB-2001
DEFINITION Halobacterium sp. NRC-1 section 103 of 170 of the complete genome.
ACCESSION AE005072 AE004437
VERSION AE005072.1 GI:10581086

KEYWORDS

SOURCE Halobacterium sp. NRC-1

ORGANISM Halobacterium sp. NRC-1

Archaea; Euryarchaeota; Halobacteria; Halobacteriales;

Halobacteriaceae; Halobacterium.

REFERENCE 1 (bases 1 to 12639)

AUTHORS Ng,W.V., Kennedy,S.P., Mahairas,G.G., Berquist,B., Pan,M.,
Shukla,H.D., Lasky,S.R., Baliga,N., Thorsson,V., Sbrogna,J.,
Swartzell,S., Weir,D., Hall,J., Dahl,T.A., Welte,R., Goo,Y.A.,
Leithauser,B., Keller,K., Cruz,R., Danson,M.J., Hough,D.W.,
Maddocks,D.G., Jablonski,P.E., Krebs,M.P., Angevine,C.M., Dale,H.,
Isenbarger,T.A., Peck,R.F., Pohlschroed,M., Spudich,J.L.,
Jung,K.-H., Alam,M., Freitas,T., Hou,S., Daniels,C.J., Dennis,P.P.,
Omer,A.D., Ebhardt,H., Lowe,T.M., Liang,P., Riley,M., Hood,L. and
DasSarma,S.

TITLE Genome sequence of Halobacterium species NRC-1

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (22), 12176-12181 (2000)

MEDLINE 20504483

PUBMED 11016950

REFERENCE 2 (bases 1 to 12639)

AUTHORS Ng,W.V., Kennedy,S.P., Mahairas,G.G., Berquist,B., Pan,M.,
Shukla,H.D., Lasky,S.R., Baliga,N., Thorsson,V., Sbrogna,J.,
Swartzell,S., Weir,D., Hall,J., Dahl,T.A., Welte,R., Goo,Y.A.,
Leithauser,B., Keller,K., Cruz,R., Danson,M.J., Hough,D.W.,
Maddocks,D.G., Jablonski,P.E., Krebs,M.P., Angevine,C.M., Dale,H.,
Isenbarger,T.A., Peck,R.F., Pohlschroed,M., Spudich,J.L.,
Jung,K.-H., Alam,M., Freitas,T., Hou,S., Daniels,C.J., Dennis,P.P.,
Omer,A.D., Ebhardt,H., Lowe,T.M., Liang,P., Riley,M., Hood,L. and
DasSarma,S.

TITLE Direct Submission

JOURNAL Submitted (14-JUL-2000) Institute for Systems Biology, 4225

FEATURES Location/Qualifiers

source

1. .12639
/organism="Halobacterium sp. NRC-1"
/mol_type="genomic DNA"
/strain="NRC-1"
/db_xref="taxon:64091"
68. .655
/gene="VNG1611C"
68. .655
/gene="VNG1611C"
/note="conserved hypothetical protein"
/codon_start=1
/transl_table=11
/product="Vng1611c"

gene

CDS

/protein_id="AAG19875.1"
/db_xref="GI:10581087"
/translation="MSAPEEEFYSEERQWLDRLREELDPEDDSARILLNLQDDV
AIAAKILNAYDSEAVSESEAVDELADIRVVLAEPAPFDDDEKMLVDGVQTSLVAVF
YSAEQYVADGVADDPVMEYVRAAVEAEAEEDLDRALGLVAAGGTRVIDGDELDMAVL
EEVEYGLVTEWVNGLDLSQSALSDEPEVIEESGE"
1896. .2705
/gene="VNG1613H"
1896. .2705
/gene="VNG1613H"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/product="Vng1613h"
/protein_id="AAG19876.1"
/db_xref="GI:10581088"
/translation="MNDDGVTGGFNTVERNEFSDAVTITDVSVLPEGGSIDLLSDDVVG
SEGRRASELYVEAEDGNQTFVDVQTTGSEYAYLGPGRMTLSTERSREEQGREESGVGL
DPVPLTTAVGGSVDVSPGEEATVEIGEFYRVPVSGPIISDPVDITDEDVRLVSVYRRGS
TSGSREFAVTGDYKPSGENFPEVTITDVSRGFGQVNVGTANANGSDSLSTISGEIVY
ASGLVRGETKSAAGESVAEQTTIPGFGPRVAVRMTVTADGNTARDVYFP"
2775. .4112
/gene="mvab"
/note="synonym: VNG1615G"
2775. .4112
/gene="mvab"
/note="mvab"
/codon_start=1
/transl_table=11
/product="3-hydroxy-3-methylglutaryl-coenzyme A synthase"
/protein_id="AAG19877.1"
/db_xref="GI:10581089"
/translation="MTSVGIDAIEIHTGKLDLAETFAPEMGDDPEKTKGLGLHAS
SFPDAYEDIVTMGANAHRLLERKGLPADVGRIDVATESAFDNSKPVSTYIAGCLEQ
VYDGFHANKGERKFACIAGTQSLDDAVNWILAGNRGRGALVIATDIALYARDSSG
EATQAGAVAMWITEDPDIVELSTEQGYSADETDFLKPNQQFPSPVDGKRSMQVYLAR
MRESLQDFERATGASHPDDYAFIPFHTPFPMVRKAALVGYRHMIRGTDEEELAADI
GRQPRRERFSDDAFEDAISEYVDGLTDTDAYREWDAAIQPTLSISSQVGNWYTGVS
HLARFSAKHAENGIDLAKRLLVASYGSAQAEIHSETVQDGFTEAFAASTIDEQI
EQRYDISFAEYEQVHDRHNHEKSVELEPFTQPDAEFVFDGWGPMNERKYTYVE"
complement (4114. .4698)
/gene="VNG1618H"
complement (4114. .4698)
/gene="VNG1618H"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/product="Vng1618h"
/protein_id="AAG19878.1"
/db_xref="GI:10581090"
/translation="MTVDTHVVAICGSLSDSGSTRALTRTVFDGATAVTDTEVTTDLI
DLRELSLPPRDGDRDAGDAPALRARVADADALVLGTPMYHGSYSAPLKNALDYCGRD
ELGTTVGLVCVAGGFPFPRQALSHLRGVCDALDADWTLPHQVGVPNAAATTIGDGEITDD
DIAARARTLGEQAVAYANVESFPELTAAGASPSQ"
complement (4761. .5570)
/gene="VNG1616C"
complement (4761. .5570)
/gene="VNG1616C"
/note="conserved hypothetical protein"
/codon_start=1
/transl_table=11
/product="Vng1616c"
/protein_id="AAG19879.1"
/db_xref="GI:10581091"
/translation="MPAPSSRRGLAVAVRAGWLDGFFPGVSQTEGMDTDDPPREGLA
VRISGEVTLSDDPGATLRKWRDFTVDSQTALADELDVSSSVISDYESGRRESPGIALV
SRLVNALLDIDERRGGTHIGQFARVLSGDFSDVLDLREYPAAVELDALYDAVDATE
VVPGRDRTVSGHTVIDSIKAIQRLSGDGFYRLYGQSTSRVLGFTNVSRGESPLVAMRV
VDPTPSAVLLHGLSPDDLWEHAPALAKRDGFSLAVTEMPVEDMLASLRQHA"
5569. .6057
/gene="VNG1617H"
5569. .6057
/gene="VNG1617H"

gene

CDS

/note="hypothetical protein"
/codon_start=1
/transl_table=11
/product="Vng1617h"
/protein_id="AAG19880.1"
/db_xref="GI:10581092"
/translation="MVDVRPYTPADEADLWELKRAFERGLGAGTGDESKAKQYADKLT
ADYRARWLDWVADCVAESPACVQVAESEDGLGVYFVLPESLAFIWDAAVLNEIFVRD
AHRGRGVADELLAAAREVVAGQSLPLDLRLVLDVDGENERAAAFYERHGFHGWELVAR
DC"
complement(6054..6626)
/gene="VNG1619H"
complement(6054..6626)
/gene="VNG1619H"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/product="Vng1619h"
/protein_id="AAG19881.1"
/db_xref="GI:10581093"
/translation="MQSVETAPTTRPATLPAAIGGITAYVVGYAVTVLWQSDAVGAAL
QSYNAVALLGGDPIPTWTAVGLFYNAHAVAFSTPAVGGSRLSRNLIIDGNAPLVLL
LVPAGLLVAGFIAGRHGDTPRADRGRLRVGVRTAGYAAVAVLGVAFRTAGTATVH
VEYALGVLVAGIGYPVFGAAGGALAAVTR"
complement(6675..7130)
/gene="VNG1621H"
complement(6675..7130)
/gene="VNG1621H"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/product="Vng1621h"
/protein_id="AAG19882.1"
/db_xref="GI:10581094"
/translation="MTDTTAVHLEPDDRDADFAPGGTGVLSFGREDAGPHSLPVSYG
YDPDTESFFRLSYGDDTDKPDPTETPLDVSVVHDEADHGWCVVVAHGRLOPTDADA
ISTTALSALRRVDIPLDFTDRPLVETEFELFRLHPDAITARADAPPTT"
7129..8634
/gene="rfcB"
/note="synonym: VNG1622G"
7129..8634
/gene="rfcB"
/note="rfcB"
/codon_start=1
/transl_table=11
/product="replication factor C large subunit"
/protein_id="AAG19883.1"
/db_xref="GI:10581095"
/translation="MPCVATERWGKTVVAGAPQHLTRTPAYGVSMVDWTEKYRPASLS
EVRGNDTARDALAEWAETWPDHREAVVHVGSPGIGKTSAAHALANDAGWDVVELNASD
QRTADVVERVAGEAARSGLTGGSGRKLVLDDADNHLGNIDRGGSAITRLVDDAP
QPIVLVANEYVEMSSLSRACREIEFRDVSRSIPLVLRDVCRRREDVTVHEDALAAIA
EQNAGDLRSVNDLQALAEQDRTLTDADVVMGERDRTEGVFDYLDVDTATHSAREALQ
AAYDVDETPDDLSSWADNVPKDYRGGELADAYEFLSNADVWLGRVATQNYAYWRYA
TDNVAAGVAAARRHDHGGWTRYGPPSYWRKLGSSRATREKRDYVARHTAFTAGCSMAT
ARNDLPLFLRLVTHCKNRELTVMAAAAYELDTEHVAFTVGSGETTNKVASIVADAEE
RRTDAAVDHSAGAFAGAVREDNTDEDSAADETTDGDDEDTGADSQRGLDEFF"
complement(8631..9464)
/gene="ccp"
/note="synonym: VNG1623G"
complement(8631..9464)
/gene="ccp"
/note="Ccp"
/codon_start=1
/transl_table=11
/product="cytochrome aa3 controlling protein"
/protein_id="AAG19884.1"
/db_xref="GI:10581096"
/translation="MARRSYQLAAVAVTLTFALLLGEFTAVSGSGATCGLQWPDCSG
QFLPFLPVNDFIEQFHRSVAMVVGFFILGTGAAYRNFDARDIKLAGVLAIVLLPLQ
VILGGTTVTFSGLVPGYQPLTQAVHHLAAMAILGTVTYALRMRDAAIGLGRVRTA
VLAALAVLPFEVAFSRGTFTVYGSWIIQIHHTFELALFTAAFTAFVWSLRNATRSA
AATGGVALLSTAQVLLGVGLFPFGPTIQLAYVALAVWIAALLVATARAHAPPTAAT

gene
CDS

gene
CDS

gene
CDS

gene A"
9595..11847
/gene="mdh"
/note="synonym: VNG1624G"

Alignment Scores:
Pred. No.: 635 Length: 12639
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 1 Gaps: 0

US-09-993-808B-2 (1-256) x AE005072 (1-12639)

QY 32 ArgSerAlaAlaAlaThrGlyValAla 41
|||||
Db 8819 CGCAGCGCGCGCCGCCACCGCGGTGTCGCG 8790
|||||

RESULT 15
AC145127_22/c
WPCOMMENT
Sequence split into 24 fragments LOCUS AC145127 Accession AC145127

| Fragment Name | Begin | End |
|---------------|---------|---------|
| AC145127_00 | 1 | 110000 |
| AC145127_01 | 100001 | 210000 |
| AC145127_02 | 200001 | 310000 |
| AC145127_03 | 300001 | 410000 |
| AC145127_04 | 400001 | 510000 |
| AC145127_05 | 500001 | 610000 |
| AC145127_06 | 600001 | 710000 |
| AC145127_07 | 700001 | 810000 |
| AC145127_08 | 800001 | 910000 |
| AC145127_09 | 900001 | 1010000 |
| AC145127_10 | 1000001 | 1110000 |
| AC145127_11 | 1100001 | 1210000 |
| AC145127_12 | 1200001 | 1310000 |
| AC145127_13 | 1300001 | 1410000 |
| AC145127_14 | 1400001 | 1510000 |
| AC145127_15 | 1500001 | 1610000 |
| AC145127_16 | 1600001 | 1710000 |
| AC145127_17 | 1700001 | 1810000 |
| AC145127_18 | 1800001 | 1910000 |
| AC145127_19 | 1900001 | 2010000 |
| AC145127_20 | 2000001 | 2110000 |
| AC145127_21 | 2100001 | 2210000 |
| AC145127_22 | 2200001 | 2310000 |
| AC145127_23 | 2300001 | 2331000 |

Continuation (23 of 24) of AC145127 from base 2200001 (AC145127 Oryza sativa (japonica cu

Alignment Scores:
Pred. No.: 3.1e+03 Length: 110000
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 8 Gaps: 0

US-09-993-808B-2 (1-256) x AC145127_22 (1-110000)

QY 28 ArgThrArgSerArgSerAlaAlaAlaThr 37
|||||
Db 31035 CGTACGCGCTCGCGCTCTGCCGCGGCGACG 31006
|||||

Search completed: October 2, 2004, 14:04:43
Job time : 3251 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2004, 12:50:51 ; Search time 2465 Seconds
(without alignments)
3101.306 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 256
Sequence: 1 MGKYMKRCRGAGAEVAE.....VRGVPLDAGGRFEWAPVWSI 256

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Word size: 1

Total number of hits satisfying chosen parameters: 55025477

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US0993808/runat_01102004_171727_13719/app_query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=oligo.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0993808 @CGN_1_1_5180 @runat_01102004_171727_13719 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:
2: em_esthum:
3: em_estin:
4: em_estmu:
5: em_estov:
6: em_estpl:
7: em_estro:
8: em_htc:
9: gb_est1:
10: gb_est2:
11: gb_htc:
12: gb_est3:
13: gb_est4:
14: gb_est5:
15: em_estfun:
16: em_estom:
17: em_gss_hum:
18: em_gss_inv:
19: em_gss_pln:
20: em_gss_vrt:
21: em_gss_fun:
22: em_gss_mam:
23: em_gss_mus:
24: em_gss_pro:
25: em_gss_rod:
26: em_gss_phg:
27: em_gss_vrl:
28: gb_gss1:

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|--------------------|
| 1 | 127 | 49.6 | 526 | 14 | CD444277 | CD444277 EL01N0438 |
| C 2 | 65 | 25.4 | 753 | 29 | CC729673 | CC729673 OGUBY35TH |
| C 3 | 65 | 25.4 | 791 | 29 | CG324885 | CG324885 OGXDN04TV |
| C 4 | 65 | 25.4 | 825 | 29 | CG265472 | CG265472 OG2CM60TH |
| C 5 | 65 | 25.4 | 890 | 29 | CG218373 | CG218373 OG1DJ10TV |
| C 6 | 65 | 25.4 | 917 | 29 | CG265481 | CG265481 OG2CM60TV |
| C 7 | 65 | 25.4 | 986 | 28 | CC002416 | CC002416 PUGKB53TB |
| C 8 | 65 | 25.4 | 987 | 28 | CC002419 | CC002419 PUGKB53TD |
| C 9 | 65 | 25.4 | 1010 | 28 | CC390569 | CC390569 PUHBT76TB |
| C 10 | 61 | 23.8 | 707 | 11 | AY110319 | AY110319 Zea mays |
| C 11 | 59 | 23.0 | 547 | 28 | BH777311 | BH777311 fzmbo13f0 |
| C 12 | 47 | 18.4 | 645 | 29 | CG136784 | CG136784 PUIDC25TB |
| C 13 | 43 | 16.8 | 488 | 28 | BH784474 | BH784474 fzmbo13f0 |
| C 14 | 43 | 16.8 | 666 | 29 | CC610974 | CC610974 OGVR67TV |
| C 15 | 43 | 16.8 | 838 | 28 | CC440801 | CC440801 PUHLN53TB |
| C 16 | 43 | 16.8 | 869 | 28 | BZ743686 | BZ743686 OGFAC93TC |
| C 17 | 42 | 16.4 | 225 | 14 | CD957606 | CD957606 SKC_144 G |
| C 18 | 41 | 16.0 | 644 | 29 | CC729680 | CC729680 OGUBY35TV |
| C 19 | 37 | 14.5 | 655 | 13 | CA104638 | CA104638 SCJFHR103 |
| C 20 | 36 | 14.1 | 659 | 13 | CA074449 | CA074449 SCEZAM108 |
| C 21 | 34 | 13.3 | 510 | 14 | CB929335 | CB929335 ABA1_41 H |
| C 22 | 29 | 11.3 | 413 | 14 | CD951880 | CD951880 SAZ_175 G |
| C 23 | 29 | 11.3 | 657 | 28 | BZ743690 | BZ743690 OGFAC93TM |
| C 24 | 28 | 10.9 | 446 | 14 | CF632445 | CF632445 zmrws48_0 |
| C 25 | 28 | 10.9 | 531 | 14 | CA297189 | CA297189 SCUTSD208 |
| C 26 | 27 | 10.5 | 563 | 13 | CA108367 | CA108367 SCJFHR104 |
| C 27 | 27 | 10.5 | 584 | 13 | CA108352 | CA108352 SCJFHR104 |
| C 28 | 27 | 10.5 | 705 | 13 | CA106521 | CA106521 SCQHR101 |
| C 29 | 27 | 10.5 | 885 | 29 | CG324875 | CG324875 OGXDN04TH |
| C 30 | 25 | 9.8 | 405 | 28 | BH879152 | BH879152 hs86e11_b |
| C 31 | 25 | 9.8 | 585 | 13 | CA104568 | CA104568 SCJFHR103 |
| C 32 | 25 | 9.8 | 661 | 14 | CA256119 | CA256119 SCJLFL418 |
| C 33 | 25 | 9.8 | 741 | 29 | CC683397 | CC683397 OGHL76TV |
| C 34 | 25 | 9.8 | 790 | 28 | CC390573 | CC390573 PUHBT76TD |
| C 35 | 25 | 9.8 | 835 | 29 | CG218361 | CG218361 OG1DJ10TH |
| C 36 | 25 | 9.8 | 887 | 29 | CG223296 | CG223296 OGWAU17TV |
| C 37 | 21 | 8.2 | 550 | 12 | BG267004 | BG267004 1000110B0 |
| C 38 | 21 | 8.2 | 568 | 9 | AI737717 | AI737717 605040C07 |
| C 39 | 21 | 8.2 | 642 | 13 | CA104689 | CA104689 SCJFHR103 |
| C 40 | 21 | 8.2 | 1197 | 11 | AY108168 | AY108168 Zea mays |
| C 41 | 17 | 6.6 | 697 | 14 | CB646543 | CB646543 OSJNEb09B |
| C 42 | 17 | 6.6 | 708 | 28 | BZ404098 | BZ404098 OGACI23TM |
| C 43 | 17 | 6.6 | 753 | 14 | CB646893 | CB646893 OSJNEb09J |
| C 44 | 17 | 6.6 | 787 | 28 | BZ404090 | BZ404090 OGACI23TC |
| C 45 | 17 | 6.6 | 795 | 29 | CG339703 | CG339703 OG1DH12TH |

ALIGNMENTS

RESULT 1
CD444277
LOCUS CD444277 526 bp mRNA linear EST 03-JUN-2003
DEFINITION EL01N0438A02.b Endosperm_4 Zea mays cDNA, mRNA sequence.
ACCESSION CD444277
VERSION CD444277.1 GI:31359920
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 526)

AUTHORS Lai, J., Dey, N., Kim, C.S., Becraft, P., Larkins, B., Linton, E. and Messing, J.
TITLE Sequencing of the maize endosperm ESTs
JOURNAL Unpublished (2002)
COMMENT Contact: Lai, Jinsheng
Dr. Joachim Messing's lab
Waksman Institute, Rutgers University
190 Frelinghuysen Rd., Piscataway, NJ 08854, USA
Tel: 732-445-3801
Fax: 732-445-5735
Email: jlai@waksman.rutgers.edu
Seq primer: T3.

FEATURES
source
1. .526
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W22"
/db_xref="taxon:4577"
/tissue type="Endosperm of 7-23DAP"
/clone_lib="Endosperm_4"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Alignment Scores:
Pred. No.: 3.99e-96 Length: 526
Score: 127.00 Matches: 127
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.61% Indels: 0
DB: 14 Gaps: 0

US-09-993-808B-2 (1-256) x CD444277 (1-526)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
|||||
Db 146 ATGGGGAAGTACATGCGCAAGTGCAGGGCGCGCAGCGCGGAGGTGCGCGCGTCGAG 205

QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal 40
|||||
Db 206 GTTACGAGTCGTCGCGTCGCGACGAGTCCAGTCCGCGCGCGGAGCCTGCTGCCGCGTGTC 265

QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaValSer 60
|||||
Db 266 GCGAAGTTCGCCCCGAGGAGGAGAGAGGCGCGCGCGGAGCCTGCTGCCGCGTGAGC 325

QY 61 AlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAla 80
|||||
Db 326 GCTGGTGGGACGCGGGAAGCTGCTACATCCACCTGCGTAGCCGATGCTGTTTCATGGCA 385

QY 81 ProProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAspGly 100
|||||
Db 386 CCGCCTCAGCGCAGCGCTCGGTTGACTCGGTTCCGACCCCGGTCGAGGCTGCTGATGGC 445

QY 101 AlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThrAla 120
|||||
Db 446 GCTGCAGACAGCAGGCGCGCGCTCGTCCGCGGCTCTCGCGGCTCTCGCTCCAGCACGGCG 505

QY 121 SerSerValAsnLeuGlyLeu 127
|||||
Db 506 TCGTCTGTGAACCTGGGCTTG 526

RESULT 2
CC729673/c

LOCUS CC729673 753 bp DNA linear GSS 23-JUN-2003

DEFINITION OGUBY35TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma0404F21,
genomic survey sequence.

ACCESSION CC729673

VERSION CC729673.1 GI:32148606

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 753)

AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGUBY35TV
Contact: Cathy Whitelaw

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source
Location/Qualifiers
1. .753
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0404F21"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Alignment Scores:
Pred. No.: 4.59e-44 Length: 753
Score: 65.00 Matches: 65
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.39% Indels: 0
DB: 29 Gaps: 0

US-09-993-808B-2 (1-256) x CC729673 (1-753)

QY 167 ArgArgGluThrThrProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAsp 186
|||||
Db 410 AGGAGAGAGACGACGCCATCGACCGGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGAT 351

QY 187 LeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeuIle 206
|||||
Db 350 CTGGCGGGCACAAAGACTGGCCCGCTGCTACCGGGCGCAACGGCGCTGCGGAGCTGATC 291

QY 207 ValProProAlaHisGluIleGlnGluPheAlaAlaAlaGluAlaAlaGlnAlaLys 226
|||||
Db 290 GTGCCGCCAGCACACGAGATCCAGAGTTCTTCGCCGCCGCGGAGCGGCCAGGCAAG 231

QY 227 ArgPheAlaSerLys 231
|||||
Db 230 CGCTTTGCTTCCAAG 216

RESULT 3
CG324885

LOCUS CG324885 791 bp DNA linear GSS 26-AUG-2003

DEFINITION OGXDN04TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma0661B08,
genomic survey sequence.

ACCESSION CG324885

VERSION CG324885.1 GI:34242151

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 791)

AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics

TITLE

JOURNAL
COMMENT

Unpublished (2002)
Other_GSSs: OGXD04TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
source

Location/Qualifiers
1. .791
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0661B08"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Alignment Scores:
Pred. No.: 4.83e-44 Length: 791
Score: 65.00 Matches: 65
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.39% Indels: 0
DB: 29 Gaps: 0

US-09-993-808B-2 (1-256) x CG324885 (1-791)

QY 167 ArgArgGluThrThrProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAsp 186
|||||
Db 196 AGGAGAGAGACGACGCCATCGAGCCGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGAT 255
QY 187 LeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeuIle 206
|||||
Db 256 CTGGCGGGGCACAAAGACTGGCCCGTCTGCTACCGCGGCAACGCCGGCTCGGAGCTGATC 315
QY 207 ValProProAlaHisGluIleGlnGluPhePheAlaAlaGluAlaAlaGlnAlaLys 226
|||||
Db 316 GTGCGCGCACACACGAGATCCAGGAGTTCTTCGCCCGCCGAGGCGGCCAGCCCAAG 375
QY 227 ArgPheAlaSerLys 231
|||||
Db 376 CGCTTTGCTTCCAAG 390

RESULT 4

CG265472/c
LOCUS CG265472 825 bp DNA linear GSS 25-AUG-2003
DEFINITION OG2CM60TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0764I23,
genomic survey sequence.
ACCESSION CG265472
VERSION CG265472.1 GI:34177613
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 825)

REFERENCE
AUTHORS

Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.

TITLE
JOURNAL

Unpublished (2002)
Other_GSSs: OG2CM60TV

COMMENT

Contact: Cathy Whitelaw
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208

FEATURES
source

Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1. .825
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0764I23"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Alignment Scores:
Pred. No.: 5.05e-44 Length: 825
Score: 65.00 Matches: 65
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.39% Indels: 0
DB: 29 Gaps: 0

US-09-993-808B-2 (1-256) x CG265472 (1-825)

QY 167 ArgArgGluThrThrProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAsp 186
|||||
Db 820 AGGAGAGAGACGACGCCATCGAGCCGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGAT 761
QY 187 LeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeuIle 206
|||||
Db 760 CTGGCGGGGCACAAAGACTGGCCCGTCTGCTACCGCGGCAACGCCGGCTCGGAGCTGATC 701
QY 207 ValProProAlaHisGluIleGlnGluPhePheAlaAlaGluAlaAlaGlnAlaLys 226
|||||
Db 700 GTGCGCGCACACACGAGATCCAGGAGTTCTTCGCCCGCCGAGGCGGCCAGCCCAAG 641
QY 227 ArgPheAlaSerLys 231
|||||
Db 640 CGCTTTGCTTCCAAG 626

RESULT 5

CG218373
LOCUS CG218373 890 bp DNA linear GSS 22-AUG-2003
DEFINITION OG1DJ10TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0736B20,
genomic survey sequence.
ACCESSION CG218373
VERSION CG218373.1 GI:34118261
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 890)

REFERENCE

AUTHORS

Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.

TITLE

JOURNAL

COMMENT

Unpublished (2002)
Other_GSSs: OG1DJ10TH

Contact: Cathy Whitelaw
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TF

Class: sheared ends.

FEATURES

source

Location/Qualifiers
1. .890
/organism="Zea mays"
/mol_type="genomic DNA"

```
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0736B20"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Alignment Scores:
Pred. No.:      5.47e-44      Length:      890
Score:          65.00         Matches:      65
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:    25.39%       Indels:        0
DB:            29           Gaps:         0

US-09-993-808B-2 (1-256) x CG218373 (1-890)

QY 167 ArgArgGluThrThrProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAsp 186
|||
Db 272 AGGCGAGAGACGACGCCATCGAGCCGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGAT 331
|||

QY 187 LeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeuIle 206
|||
Db 332 CTGGCGGGGCACAAAGACTGGCCCGTCCGTACCGCGGCAACGCCGGTGGAGCTGATC 391
|||

QY 207 ValProProAlaHisGluIleGlnGluPhePheAlaAlaGluAlaGlnAlaLys 226
|||
Db 392 GTGCCGCCACACACGAGATCCAGGAGTCTTCGCCCGCGCGAGGCGGCCCAAG 451
|||

QY 227 ArgPheAlaSerLys 231
|||
Db 452 CGCTTTGCTTCCAAG 466
|||

RESULT 6
CG265481
LOCUS
DEFINITION
  OG2CM60TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0764I23,
  genomic survey sequence.
ACCESSION
  CG265481
VERSION
  CG265481.1 GI:34177622
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
  1 (bases 1 to 917)
REFERENCE
  Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
  Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
  Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  Consortium for Maize Genomics
  Unpublished (2002)
  Other_GSSs: OG2CM60TH
  Contact: Cathy Whitelaw
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@tigr.org
  Seq primer: TF
  Class: sheared ends.
  Location/Qualifiers
    1..917
      /organism="Zea mays"
      /mol_type="genomic DNA"
      /strain="B73"
      /db_xref="taxon:4577"
      /clone="ZMMBMA0764I23"
      /clone_lib="ZM_0.7_1.5_KB"
      /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
      methylation filtered genomic DNA library"

ORIGIN

Alignment Scores:
Pred. No.:      5.47e-44      Length:      890
Score:          65.00         Matches:      65
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:    25.39%       Indels:        0
DB:            29           Gaps:         0

US-09-993-808B-2 (1-256) x CG218373 (1-890)

QY 167 ArgArgGluThrThrProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAsp 186
|||
Db 272 AGGCGAGAGACGACGCCATCGAGCCGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGAT 331
|||

QY 187 LeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeuIle 206
|||
Db 332 CTGGCGGGGCACAAAGACTGGCCCGTCCGTACCGCGGCAACGCCGGTGGAGCTGATC 391
|||

QY 207 ValProProAlaHisGluIleGlnGluPhePheAlaAlaGluAlaGlnAlaLys 226
|||
Db 392 GTGCCGCCACACACGAGATCCAGGAGTCTTCGCCCGCGCGAGGCGGCCCAAG 451
|||

QY 227 ArgPheAlaSerLys 231
|||
Db 452 CGCTTTGCTTCCAAG 466
|||

RESULT 6
CG265481
LOCUS
DEFINITION
  OG2CM60TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0764I23,
  genomic survey sequence.
ACCESSION
  CG265481
VERSION
  CG265481.1 GI:34177622
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
  1 (bases 1 to 917)
REFERENCE
  Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
  Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
  Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
  Consortium for Maize Genomics
  Unpublished (2002)
  Other_GSSs: OG2CM60TH
  Contact: Cathy Whitelaw
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@tigr.org
  Seq primer: TF
  Class: sheared ends.
  Location/Qualifiers
    1..917
      /organism="Zea mays"
      /mol_type="genomic DNA"
      /strain="B73"
      /db_xref="taxon:4577"
      /clone="ZMMBMA0764I23"
      /clone_lib="ZM_0.7_1.5_KB"
      /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
      methylation filtered genomic DNA library"

ORIGIN
```

```
Alignment Scores:
Pred. No.:      5.64e-44      Length:      917
Score:          65.00         Matches:      65
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:    25.39%       Indels:        0
DB:            29           Gaps:         0

US-09-993-808B-2 (1-256) x CG265481 (1-917)

QY 167 ArgArgGluThrThrProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAsp 186
|||
Db 262 AGGAGAGAGACGACGCCATCGAGCCGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGAT 321
|||

QY 187 LeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeuIle 206
|||
Db 322 CTGGCGGGGCACAAAGACTGGCCCGTCCGTACCGCGGCAACGCCGGTGGAGCTGATC 381
|||

QY 207 ValProProAlaHisGluIleGlnGluPhePheAlaAlaGluAlaGlnAlaLys 226
|||
Db 382 GTGCCGCCACACACGAGATCCAGGAGTCTTCGCCCGCGCGAGGCGGCCCAAG 441
|||

QY 227 ArgPheAlaSerLys 231
|||
Db 442 CGCTTTGCTTCCAAG 456
|||
```

```
RESULT 7
CC002416
LOCUS
DEFINITION
  PUGKB53TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa402I10,
  genomic survey sequence.
ACCESSION
  CC002416
VERSION
  CC002416.1 GI:29380976
KEYWORDS
  GSS.
SOURCE
  Zea mays
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
    clade; Panicoideae; Andropogoneae; Zea.
  1 (bases 1 to 986)
REFERENCE
  Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
  Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
  Bennetzen,J.
  Maize Genomics Consortium
  Unpublished (2003)
  Other_GSSs: PUGKB53TD
  Contact: Cathy Whitelaw
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@tigr.org
  Seq primer: TR
  Class: sheared ends.
  Location/Qualifiers
    1..986
      /organism="Zea mays"
      /mol_type="genomic DNA"
      /strain="B73"
      /db_xref="taxon:4577"
      /clone="ZMMBTa402I10"
      /clone_lib="ZM_0.6_1.0_KB"
      /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
      Cot selected genomic DNA library"
```

```
FEATURES
source
  1..986
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone="ZMMBTa402I10"
    /clone_lib="ZM_0.6_1.0_KB"
    /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
    Cot selected genomic DNA library"

ORIGIN

Alignment Scores:
Pred. No.:      6.09e-44      Length:      986
Score:          65.00         Matches:      65
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:    25.39%       Indels:        0
```

```
DB:                28                Gaps:                0
US-09-993-808B-2 (1-256) x CC002416 (1-986)
QY 167 ArgArgGluThrThrProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAsp 186
Db 346 AGGCGAGAGACGACGCCATCGAGCCGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGAT 405
QY 187 LeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeu 206
Db 406 CTGGCGGGGACAAAGACTGGCCCGTACCGGGGCAACGCCGGGTGCGGAGCTGATC 465
QY 207 ValProProAlaHisGluIleGlnGluPheAlaAlaGluAlaGlnAlaLys 226
Db 466 GTGCCGCCAGCACACGAGATCCAGGAGTTCTTTGCCCGCCGAGCGGCCAGGCCAAG 525
QY 227 ArgPheAlaSerLys 231
Db 526 CGCTTTGCTTCCAAG 540

RESULT 8
CC002419/c
LOCUS CC002419 987 bp DNA linear GSS 31-MAR-2003
DEFINITION PUGKB53TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa402I10,
genomic survey sequence.
ACCESSION CC002419
VERSION CC002419.1 GI:29380979
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 987)
REFERENCE Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
AUTHORS Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: PUGKB53TB
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
    source
        1..987
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMMBTa402I10"
            /clone_lib="ZM_0.6_1.0_KB"
            /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
CoT selected genomic DNA library"
ORIGIN
Alignment Scores: 6.1e-44 Length: 987
Pred. No.: 65.00 Matches: 65
Score: 100.00%
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 25.39%
DB: 28 Gaps: 0

US-09-993-808B-2 (1-256) x CC002419 (1-987)
QY 167 ArgArgGluThrThrProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAsp 186
Db 640 AGGCGAGAGACGACGCCATCGAGCCGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGAT 581
```

```
QY 187 LeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeu 206
Db 580 CTGGCGGGGACAAAGACTGGCCCGTACCGGGGCAACGCCGGGTGCGGAGCTGATC 521
QY 207 ValProProAlaHisGluIleGlnGluPheAlaAlaGluAlaGlnAlaLys 226
Db 520 GTGCCGCCAGCACACGAGATCCAGGAGTTCTTTGCCCGCCGAGCGGCCAGGCCAAG 461
QY 227 ArgPheAlaSerLys 231
Db 460 CGCTTTGCTTCCAAG 446

RESULT 9
CC390569
LOCUS CC390569 1010 bp DNA linear GSS 19-MAY-2003
DEFINITION PUHBT76TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa415M08,
genomic survey sequence.
ACCESSION CC390569
VERSION CC390569.1 GI:30870659
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 1010)
REFERENCE Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
AUTHORS Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: PUHBT76TD
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
    Location/Qualifiers
        1..1010
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMMBTa415M08"
            /clone_lib="ZM_0.6_1.0_KB"
            /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
CoT selected genomic DNA library"
ORIGIN
Alignment Scores: 6.24e-44 Length: 1010
Pred. No.: 65.00 Matches: 65
Score: 100.00%
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 25.39%
DB: 28 Gaps: 0

US-09-993-808B-2 (1-256) x CC390569 (1-1010)
QY 167 ArgArgGluThrThrProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAsp 186
Db 177 AGGCGAGAGACGACGCCATCGAGCCGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGAT 236
QY 187 LeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeu 206
Db 237 CTGGCGGGGACAAAGACTGGCCCGTACCGGGGCAACGCCGGGTGCGGAGCTGATC 296
QY 207 ValProProAlaHisGluIleGlnGluPheAlaAlaGluAlaGlnAlaLys 226
```


Db 297 GTGCGCCAGCACACGAGATCCAGGAGTTCTTCGCCGCCGCGAGCGGCCAGGCCAG 356

QY 227 ArgpheaLaserLys 231
|||||

Db 357 CGCTTGTCTTCCAAG 371

RESULT 10
AY110319

LOCUS
DEFINITION
Zea mays CL528_-1 mRNA sequence.
ACCESSION
AY110319
VERSION
AY110319.1 GI:21214635
KEYWORDS
HTC.

SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1 (bases 1 to 707)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 707)
Coe,E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB and may be found by BLAST
searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.

FEATURES
source
Location/Qualifiers
1..707
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:632312"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN
Alignment Scores:
Pred. No.: 9.58e-41 Length: 707
Score: 61.00 Matches: 61
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 23.83% Indels: 0
DB: 11 Gaps: 0

US-09-993-808B-2 (1-256) x AY110319 (1-707)

QY 156 AlaSerAsnSerGlySerGlyProAspArgGluArgGluThrThrProSerSerArg 175
|||||

Db 82 GCGAGCAACTCCGGAGCGGCCAGACCGGAGAGAGAGACGACGCGCATCGAGCCGG 141

QY 176 AlaHisGlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSer 195
|||||

Db 142 GCGACGCGGAGCTCAGCGATCTGGAGTCGATCTGGCGGGGCACAGACTGGCCCGTCG 201

QY 196 LeuProAlaAlaThrProAlaAlaGluLeuLeuValProProAlaAlaHisGluLeuGlu 215
|||||

Db 202 CTACCGGGCGGCAACGCGGCTGCGGAGCTGATCGTCCCGCCAGCACACGAGATCCAGGAG 261

QY 216 Phe 216
|||||

Db 262 TTC 264

RESULT 11
BH777311/c

LOCUS
DEFINITION
fzmb013f023f06f0 fzmb filtered library Zea mays genomic clone
fzmb013f023f06 5', genomic survey sequence.
ACCESSION
BH777311
VERSION
BH777311.1 GI:19779747
KEYWORDS
GSS.

SOURCE
Zea mays
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1 (bases 1 to 547)
Budiman,M.A., Freese,R.G., Bedell,J.A., Nunberg,A.N. and Lakey,N.D.
GeneThresher methylation filtered genomic sequences from maize
Unpublished (2002)
Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: fzmb013f023 row: f column: 06
Seq primer: M13 forward
Class: shotgun
High quality sequence stop: 547.

FEATURES
source
Location/Qualifiers
1..547
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="Mol7"
/db_xref="taxon:4577"
/clone="fzmb013f023f06"
/clone_lib="fzmb filtered library"
/note="Organ: leaf; Vector: pBCSK(-); Site:1: HincII; DNA
prepared from purified nuclei was randomly sheared,
end-repaired, size fractionated to enrich for the 0.5 to
5 kb fraction, ligated into HincII-digested pBCSK(-)
vector and electroporated into E. coli cells."

ORIGIN
Alignment Scores:
Pred. No.: 3.46e-39 Length: 547
Score: 59.00 Matches: 59
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 23.05% Indels: 0
DB: 28 Gaps: 0

US-09-993-808B-2 (1-256) x BH777311 (1-547)

QY 167 ArgArgGluThrThrProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSer 186
|||||

Db 294 AGGCGAGAGACGACGCGCATCGAGCGGGCGCACGGCGAGCTCAGCGATCTGGAGTCGGAT 235

QY 187 LeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeu 206
|||||

Db 234 CTGGCGGGGCACAGACTGGCCCGTCGCTACCGCGGCAACGCCGCTCGGAGCTGATC 175

QY 207 ValProProAlaHisGluLeuGlnGluPheAlaAlaAlaGluAlaGlnAla 225
|||||

Db 174 GTGCCGCCAGCACACGAGATCCAGGAGTCTTCGCCCGCCGAGCGGCCAGGCC 118

RESULT 12
CG136784

LOCUS
DEFINITION
CG136784 645 bp DNA linear GSS 21-AUG-2003
PUIDC25TB_ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa0556E01,

genomic survey sequence.
ACCESSION CG136784
VERSION CG136784.1 GI:34026526
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 645)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: PUIDC25TD
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..645
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa0556E01"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high CoT selected genomic DNA library"
ORIGIN
Alignment Scores:
Pred. No.: 4.56e-29 Length: 645
Score: 47.00 Matches: 47
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.36% Indels: 0
DB: 29 Gaps: 0
US-09-993-808B-2 (1-256) x CG136784 (1-645)
QY 167 ArgArgGluThrThrProSerSerArgAlaHisGlyGluLeuSerAspLeuGluSerAsp 186
|||||
Db 503 AGGCGAGACGACGCCATCGAGCCGGCGCACGGGAGCTCAGCGATCTGGAGTCGGAT 562
187 LeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThrProAlaAlaGluLeuIle 206
|||||
Db 563 CTGGCGGGGCACAAGACTGGCCCCGTGCTACCGGGCAACCGCGGTGCGGAGTGATC 622
QY 207 ValProProAlaHisGluIle 213
|||||
Db 623 GTGCCGCCAGCACACGAGATC 643
RESULT 13
BH784474
LOCUS fzm013f023f06k0 fzm filtered library Zea mays genomic clone
DEFINITION fzm013f023f06 5', genomic survey sequence.
ACCESSION BH784474
VERSION BH784474.1 GI:19788171
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 488)
AUTHORS Budiman,M.A., Freese,R.G., Bedell,J.A., Nunberg,A.N. and Lakey,N.D.

GeneThresher methylation filtered genomic sequences from maize
Unpublished (2002)
Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: fzm013f023 row: f column: 06
Seq primer: SK reverse
Class: shotgun
High quality sequence stop: 488.
Location/Qualifiers
1..488
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="Mol7"
/db_xref="taxon:4577"
/clone="fzm013f023f06"
/clone_lib="fzm filtered library"
/note="Organ: leaf; Vector: pBCSK(-); Site 1: HincII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5 kb fraction, ligated into HincII-digested pBCSK(-) vector and electroporated into E. coli cells."
ORIGIN
Alignment Scores:
Pred. No.: 7.59e-26 Length: 488
Score: 43.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.80% Indels: 0
DB: 28 Gaps: 0
US-09-993-808B-2 (1-256) x BH784474 (1-488)
QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
|||||
Db 323 ATGGGGAAGTACATGCGCAAGTGCAGGGGGCGCGAGGCGCGGCGGTCGCGTGCAG 382
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaThrGlyGlyVal 40
|||||
Db 383 GTTACGCGAGTCTCGGCGTCCGCGACGAGGTCCAGGTCCGCGGCGGCGGCGGCGGTGC 442
QY 41 AlaLysVal 43
|||||
Db 443 GCGAAGGTC 451
RESULT 14
CC610974
LOCUS OGVFR67TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma0551L14,
DEFINITION genomic survey sequence.
ACCESSION CC610974
VERSION CC610974.1 GI:31972395
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 666)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfsing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGVFR67TH
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843

Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
source

Location/Qualifiers
1..666
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0551L14"
/clone_lib="ZM 0.7 1.5 KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Alignment Scores:
Pred. No.: 1.05e-25 Length: 666
Score: 43.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.80% Indels: 0
DB: 29 Gaps: 0

US-09-993-808B-2 (1-256) x CC610974 (1-666)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
|||||
Db 467 ATGGGGAAGTACATGCGCAAGTGCAGGGGGCCGAGGCGGGAGGTCCGCCGTCGAG 408
|||||

QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal 40
|||||
Db 407 GTTACGCGAGTGTGTCGGCGTCCGCACGAGGTCCAGGTCCGCGGGCGGCGGCGGTGTC 348
|||||

QY 41 AlaLysVal 43
|||||
Db 347 GCGAAGGTC 339

RESULT 15
CC440801/c
LOCUS
DEFINITION
CC440801 PUHLN53TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa478J10,
genomic survey sequence.
CC440801
VERSION CC440801.1 GI:30942238
KEYWORDS GSS.
SOURCE
ORGANISM
Zea mays

REFERENCE
AUTHORS
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 838)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE
JOURNAL
COMMENT
Maize Genomics Consortium
Unpublished (2003)
Other_GSSs: PUHLN53TD
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source
Location/Qualifiers
1..838
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa478J10"

/clone_lib="ZM 0.6 1.0 KB"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
CoT selected genomic DNA library"

ORIGIN

Alignment Scores:
Pred. No.: 1.34e-25 Length: 838
Score: 43.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.80% Indels: 0
DB: 28 Gaps: 0

US-09-993-808B-2 (1-256) x CC440801 (1-838)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
|||||
Db 390 ATGGGAAAGTACATGCGCAAGTGCAGGGGGCCGAGGCGGGAGGTCCGCCGTCGAG 331
|||||

QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGlyVal 40
|||||
Db 330 GTTACGCGAGTGTGTCGGCGTCCGCACGAGGTCCAGGTCCGCGGGCGGCGGCGGTGTC 271
|||||

QY 41 AlaLysVal 43
|||||
Db 270 GCGAAGGTC 262

Search completed: October 2, 2004, 14:43:24
Job time : 2470 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2004, 11:51:15 ; Search time 366 Seconds
(without alignments)
2971.417 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 256
Sequence: 1 MGKYMRCRGAAGAEVAE.....VRGVPLDAGRFEWAPVWSI 256

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Word size: 1 6745146
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09993808/runat_01102004_171726_13699/app_query.fasta_1.455
-DB=N Geneseq_29Jan04 -QFMT=fastap -SUFFIX=oligo.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09993808 @CGN 1 1 708 @runat_01102004_171726_13699 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04.*
1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002s.*
7: geneseqn2003as.*
8: geneseqn2003bs.*
9: geneseqn2003cs.*
10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 256 | 100.0 | 1372 | 7 ABV74603 | Abv74603 Maize CKI |
| 2 | 101 | 39.5 | 723 | 3 AAA95281 | Aaa95281 Corn cycl |
| 3 | 21 | 8.2 | 568 | 6 ABK93954 | Abk93954 Corn zmIC |
| 4 | 21 | 8.2 | 572 | 3 AAN02390 | Aan02390 Corn cycl |
| 5 | 21 | 8.2 | 572 | 3 AAA95276 | Aaa95276 Corn cycl |
| 6 | 21 | 8.2 | 639 | 3 AAA95280 | Aaa95280 Corn cycl |
| 7 | 17 | 6.6 | 841 | 7 ABV74605 | Abv74605 Maize CKI |
| 8 | 17 | 6.6 | 1242 | 6 ABK93958 | Abk93958 DNA encod |

| | | | | | | |
|----|----|-----|------|---|-----------|---------------------|
| 9 | 10 | 3.9 | 681 | 6 | ABK933981 | Abk933981 Full leng |
| 10 | 10 | 3.9 | 1073 | 6 | ABK933982 | Abk933982 Rice OsIC |
| 11 | 10 | 3.9 | 1539 | 7 | ACA42627 | Aca42627 Prokaryot |
| 12 | 10 | 3.9 | 3525 | 7 | ADA71170 | Ada71170 Rice gene |
| 13 | 9 | 3.5 | 204 | 3 | AAA95286 | Aaa95286 Rice cycl |
| 14 | 9 | 3.5 | 204 | 3 | AAN02400 | Aan02400 Rice Cycl |
| 15 | 9 | 3.5 | 334 | 6 | ABL81995 | AbL81995 Human ova |
| 16 | 9 | 3.5 | 392 | 4 | AAK56036 | Aak56036 Human imm |
| 17 | 9 | 3.5 | 414 | 8 | ACH30301 | Ach30301 Human tes |
| 18 | 9 | 3.5 | 436 | 2 | AAV38676 | Aav38676 Mus muscu |
| 19 | 9 | 3.5 | 484 | 7 | ACA35095 | Aca35095 Prokaryot |
| 20 | 9 | 3.5 | 493 | 6 | ABK93950 | Abk93950 Rice OsIC |
| 21 | 9 | 3.5 | 499 | 8 | ACL13592 | ACL13592 DNA clone |
| 22 | 9 | 3.5 | 500 | 8 | ACL25929 | ACL25929 DNA clone |
| 23 | 9 | 3.5 | 501 | 8 | ACL25927 | ACL25927 DNA clone |
| 24 | 9 | 3.5 | 529 | 8 | ACL25925 | ACL25925 DNA clone |
| 25 | 9 | 3.5 | 530 | 8 | ACL25926 | ACL25926 DNA clone |
| 26 | 9 | 3.5 | 571 | 8 | ACL13581 | ACL13581 DNA clone |
| 27 | 9 | 3.5 | 571 | 8 | ACL13580 | ACL13580 DNA clone |
| 28 | 9 | 3.5 | 578 | 8 | ACL13594 | ACL13594 DNA clone |
| 29 | 9 | 3.5 | 597 | 8 | ACL25923 | ACL25923 DNA clone |
| 30 | 9 | 3.5 | 630 | 8 | ACL13582 | ACL13582 DNA clone |
| 31 | 9 | 3.5 | 631 | 8 | ACL13590 | ACL13590 DNA clone |
| 32 | 9 | 3.5 | 632 | 8 | ACL13589 | ACL13589 DNA clone |
| 33 | 9 | 3.5 | 634 | 8 | ACL13591 | ACL13591 DNA clone |
| 34 | 9 | 3.5 | 634 | 8 | ACL13595 | ACL13595 DNA clone |
| 35 | 9 | 3.5 | 634 | 8 | ACL13584 | ACL13584 DNA clone |
| 36 | 9 | 3.5 | 635 | 8 | ACL25924 | ACL25924 DNA clone |
| 37 | 9 | 3.5 | 646 | 8 | ACL13593 | ACL13593 DNA clone |
| 38 | 9 | 3.5 | 663 | 8 | ACL25922 | ACL25922 DNA clone |
| 39 | 9 | 3.5 | 663 | 8 | ACL13579 | ACL13579 DNA clone |
| 40 | 9 | 3.5 | 667 | 8 | ACL13596 | ACL13596 DNA clone |
| 41 | 9 | 3.5 | 675 | 8 | ACL25928 | ACL25928 DNA clone |
| 42 | 9 | 3.5 | 677 | 8 | ACL13583 | ACL13583 DNA clone |
| 43 | 9 | 3.5 | 677 | 8 | ACL13577 | ACL13577 DNA clone |
| 44 | 9 | 3.5 | 822 | 3 | AAA45372 | Aaa45372 Human sec |
| 45 | 9 | 3.5 | 975 | 7 | ABZ66676 | Abz66676 Orthosomy |

ALIGNMENTS

RESULT 1
ABV74603
ID ABV74603 standard; DNA; 1372 BP.
XX
AC ABV74603;
XX
XX
DT 21-FEB-2003 (first entry)
XX
XX
DE Maize CKI_B coding sequence.
XX
KW Maize; cyclin-dependent kinase inhibitor; CKI; CKI_B; plant; corn;
KW crop yield; root size; plant growth; tassel size; ear size;
KW male sterility; endoreduplication; gene; ds.
XX
OS Zea mays.
XX
XX
FH Key Location/Qualifiers
CDS 134..904
FT /*tag= a
FT /product= "CKI_B"
XX
PN WO200281623-A2.
XX
PD 17-OCT-2002.
XX
XX
PF 06-NOV-2001; 2001WO-US044038.
XX
XX
PR 07-NOV-2000; 2000US-0246349P.
XX
XX
PA (PION-) PIONEER HI-BRED INT INC.
PA (ARIZ-) ARIZONA BOARD OF REGENTS.
XX

PI Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;
XX WPI; 2003-058511/05.
DR P-PSDB; ABB98757.
XX Novel cyclin-dependent kinase polynucleotides and their encoded proteins,
PT involved in cell cycle regulation, and useful for altering cell cycle
PT protein content, cell cycle progression, cell number and composition of
PT plants.
XX Claim 1; Page 64-65; 69pp; English.
PS The present sequence is the coding sequence (I) for maize cyclin-
XX dependent kinase inhibitor (CKI), CKI B. (I) is useful for modulating the
CC activity of cyclin-dependent kinase (CDK) in a plant such as a corn,
CC soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice,
CC barley, oil-seed Brassica and millet. Modulating the activity of CDK,
CC preferably modulating downward is useful for providing differential
CC growth in a plant, especially a positive growth advantage and modulating
CC CDK activity upward is useful for increasing crop yield, root size, plant
CC growth, tassel size and/or ear size. Modulating CDK activity is also
CC useful for conferring male sterility and for improving transformation
CC frequencies by increasing the number of cells in cell division. CDK
CC activity can also be modulated for modulating endoreduplication in the
CC endosperm of corn, sorghum, wheat, rice, barley, and millet, where the
CC promoter used is an endosperm-preferred promoter. The cell numbers are
CC modulated in one or more tissues of a plant, comprising root, seed,
CC tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves,
CC stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm,
CC pericarp, meristematic or leaf cells. (I) is also useful for identifying
CC maize CKI interacting proteins, by adducting the nucleic acid sequence to
CC a second nucleic acid sequence encoding a DNA-binding domain
XX
SQ Sequence 1372 BP; 292 A; 396 C; 442 G; 241 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 4.01e-226 Length: 1372
Score: 256.00 Matches: 256
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-09-993-808B-2 (1-256) x ABV74603 (1-1372)

QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaValGlu 20
DB 134 ATGGGAAGTACATCGCAAGTCCAGGGCGCCGAGGCGGAGGTGCGCCGCGTCCGAG 193
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyVal 40
DB 194 GTTACGAGGTGCTCGGCGTCCGACGAGGTCCAGGTCCGCGGGGAGCTGCTCCGCGGTGTC 253
QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaValSer 60
DB 254 GCGAAGTGCCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 313
QY 61 AlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAla 80
DB 314 GCTGGTGGGACGGCGGAGGAGTCTACATCCACCTGCTAGCGATGCTGTTTCATGGCA 373
QY 81 ProProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAspGly 100
DB 374 CCGCTCAGCCGCGAGCCGCTCGGTTCGACTCCGACCCCGGCTGAGGCTGCTGATGGC 433
QY 101 AlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThrAla 120
DB 434 GCTGCAGGACAGCAGGGCGGCGGCTCGCGCGGGGCTCTCGCGTGTCTCCAGCACGGCG 493
QY 121 SerSerValAsnLeuGlyLeuGlyGlnArgGlySerHisThrCysArgSerTyrAsp 140
DB 494 TCGTCGGTGAACCTGGGCTTGGGGGTTCAGCGCGGGAGCCACACCTGCCGCTCTACGAC 553

QY 141 AlaAlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGly 160
DB 554 GCTGCAGAGGCTGCGGGGATCACGTCCTGTGTGATGTCTCGGGGGGAGCAACTCCGGG 613
QY 161 SerGlyProAspArgGluArgArgGluThrThrProSerSerArgAlaHisGlyGluLeu 180
DB 614 AGCGGCCAGACCGGAGAGGCGAGAGACGACGACCATCGAGCCGGCGGCGAGGCTC 673
QY 181 SerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThr 200
DB 674 AGCGATCTGGAGTCCGATCTGGCGGGGCACAACTGCGCTCGTACCGCGGCAACG 733
QY 201 ProAlaAlaGluLeuValProProAlaHisGluIleGlnGluPhePheAlaAlaAla 220
DB 734 CCGGCTGCGGAGCTGATCGTCCGCCAGCACACGAGATCCAGGAGTTCTTCGCGCGCGC 793
QY 221 GluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyVal 240
DB 794 GAGCGGGCCAGGCCAAGCGCTTGTCTCCAAAGTACAACTTCGACTTCGTCCGCGCGGTG 853
QY 241 ProLeuAspAlaGlyGlyArgPheGluTrpAlaProValValSerIle 256
DB 854 CCCCTCGACGCCGCGCGGTTCGAGTGGGCGCGCGGTGTCAGCATC 901

RESULT 2

AAA95281
ID AAA95281 standard; cDNA; 723 BP.

XX AAA95281;

AC AAA95281;

XX 17-JAN-2001 (first entry)

XX Corn cyclin-dependent kinase inhibitor coding sequence #3.

DE Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;

XX cell growth; herbicide; ss.

OS Zea mays.

XX Key Location/Qualifiers

FT CDS 3..380

FT /*tag= a

FT /product= "CDKI"

FT /partial

XX WO200060087-A2.

XX 12-OCT-2000.

XX 06-APR-2000; 2000WO-US009106.

XX 07-APR-1999; 99US-0128192P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Klein TM, Weng Z, Cahoon RE;

XX WPI; 2000-679375/66.

XX P-PSDB; AAB26250.

XX Cyclin dependent kinase inhibitor sequences, useful for identifying

XX herbicides and plant growth inhibitors.

XX Claim 2; Page 43-44; 58pp; English.

XX The present sequence is the coding sequence for the corn cyclin-dependent
CC kinase inhibitor (CDKI). It was isolated by searching a contig comprising
CC cDNA from corn roots and ear leaf sheath for sequences similar to those
CC encoding the CDKI from *Chenopodium rubrum*, *Caenorhabditis elegans* and
CC *Arabidopsis thaliana*. CDKI is involved in the cell cycle, and may promote
CC or inhibit cell division and growth. The coding sequence, and the protein
CC it encodes are useful in the production of transgenic plants which
CC produce increased or decreased amounts of the CDKI protein, in the

CC identification of herbicides, in genetic and physical mapping and in the
CC isolation of the CDKI gene in other organisms
XX
SQ Sequence 723 BP; 160 A; 202 C; 210 G; 151 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.59e-83 Length: 723
Score: 101.00 Matches: 101
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 39.45% Indels: 0
DB: 3 Gaps: 0

US-09-993-808B-2 (1-256) x AAA95281 (1-723)

QY 156 AlaSerAsnSerGlySerGlyProAspArgGluArgArgGluThrThrProSerSerArg 175
Db 75 GCGAGCAACTCCGGGAGCGGCCAGACCGGAGAGAGAGACGACGCGCATCGAGCCGG 134
QY 176 AlaHisGlyGluLeuSerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSer 195
Db 135 GCGACGGCGAGCTCAGCGATCTGGAGTCGGATCTGGCGGGGACAAAGACTGGCCCGTCG 194
QY 196 LeuProAlaAlaThrProAlaAlaGluLeuValProProAlaHisGluIleGlnGlu 215
Db 195 CTACCGGGGCAACGCCGCTGCGAGCTGATCGTGCCGCGACACGAGATCCAGGAG 254
QY 216 PhePheAlaAlaAlaGluAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAsp 235
Db 255 TTCTTCGCGCGCCGCGAGCGGCCAGGCCAAGCGCTTGTCTTCCAAAGTACAACTTCGAC 314
QY 236 PheValArgGlyValProLeuAspAlaGlyGlyArgPheGluTrpAlaProValValSer 255
Db 315 TTCGTCCGCGCGGTGCCCTCGACGCGCGCGCGGTTCGAGTGGCGCGCGTGTGTGACG 374
QY 256 Ile 256
Db 375 ATC 377

RESULT 3
ABK93954
ID ABK93954 standard; DNA; 568 BP.
XX
AC ABK93954;
XX

DT 07-AUG-2003 (revised)
DT 27-AUG-2002 (first entry)
XX

DE Corn zmICK1 EST DNA sequence.
XX

KW Plant; ds; gene; inhibitor of cyclin dependent kinase; ICK.
XX

OS Zea mays.
XX

PN WO200228893-A2.
XX

PD 11-APR-2002.
XX

PF 29-JUN-2001; 2001WO-IB001492.
XX

PR 14-JUL-2000; 2000US-0218471P.
XX

PR 13-OCT-2000; 2000US-0241219P.
XX

PA (CROP-) CROPDESIGN NV.
XX

PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
XX

PI Hatzfeld Y;
XX

XX WPI; 2002-471311/50.
DR

DR P-PSDB; ABG65674.
XX

PT Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
to screen substrates, drugs or compounds which modulate ICK activity and
PT

PT treat disorders characterized by an insufficient or excessive production
PT of ICK inhibitors.
XX

PS Claim 41; Page 125; 141pp; English.
XX

CC This invention relates to the DNA and protein sequences of novel isolated
CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
CC the invention may be used for treating disorders characterised by
CC insufficient or excessive production of an ICK inhibitor. The protein of
CC the invention may also be used to screen for naturally-occurring ICK
CC substrates, drugs or compounds which modulate ICK activity, as well as to
CC treat disorders characterised by insufficient or excessive production of
CC ICK protein, forms which have decreased or aberrant activity compared to
CC ICK wild type protein. The present sequence represents a cDNA molecule
CC encoding an inhibitor of cyclin dependent kinase (ICK) protein of the
CC invention. (Updated on 07-AUG-2003 to correct OS field.)
XX

SQ Sequence 568 BP; 139 A; 149 C; 177 G; 102 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 6.85e-10 Length: 568
Score: 21.00 Matches: 21
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.20% Indels: 0
DB: 6 Gaps: 0

US-09-993-808B-2 (1-256) x ABK93954 (1-568)

QY 225 AlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyValProLeuAspAla 244
Db 257 GCCAAACGCTTTGCTTCCAAAGTACAACTTCGACTTCGTCGCGCGGTGCCCTCGACGCC 316
QY 245 Gly 245
Db 317 GGC 319

RESULT 4
AAN02390
ID AAN02390 standard; DNA; 572 BP.
XX
AC AAN02390;
XX

DT 01-NOV-2001 (first entry)
XX

DE Corn Cyclin dependent kinase inhibitor (CDKI) clone csiln.pk0050.e6.
XX

KW Cyclin dependent kinase inhibitor; CDKI; herbicide; cell cycle; corn;
KW plant growth inhibitor; ds.
XX

OS Zea mays.
XX

XX Location/Qualifiers
FH Key 208.369
CDS /*tag= a

FT /product= "CDKI fragment"
FT /partial

FT /note= "No start or stop codon given"
XX

XX WO200060087-A2.
PN

XX 12-OCT-2000.
PD

XX 06-APR-2000; 2000WO-US009106.
PF

XX 07-APR-1999; 99US-0128192P.
PR

XX (DUPO) DU PONT DE NEMOURS & CO E I.
PA

XX Klein TM, Weng Z, Cahoon RE;
PI

XX WPI; 2000-679375/66.
XX

DR P-PSDB; AAP01940.
DR

XX Cyclin dependent kinase inhibitor sequences, useful for identifying
PT herbicides and plant growth inhibitors.
PT
XX
PS Claim 2; Page 39-40; 58pp; English.
XX
CC The invention describes a novel isolated polynucleotide comprising a
CC nucleotide sequence encoding one of 17 specific cyclin dependent kinase
CC inhibitor (CDKI) polypeptides, cell cycle regulators involved in control
CC of cell division, growth and death. The nucleotide sequences can be used
CC in a vector to transform a host cell to produce the CDKI polypeptide.
CC They can also be used in methods for selecting and obtaining a nucleic
CC acid sequence that encodes CDKI or affects the level of CDKI expression.
CC The encoded protein can be used in a method for evaluating a compound for
CC its ability to inhibit the activity of a CDKI. The inhibitors can be used
CC as herbicides. They can also be used to inhibit plant growth. The
CC polynucleotide sequences can be used in gene mapping and as genetic
CC markers. The sequence encodes the corn CDKI clone csiln.pk0050.e6 as
CC described in the method of the invention
XX
SQ Sequence 572 BP; 140 A; 156 C; 176 G; 93 T; 0 U; 7 Other;

Alignment Scores:
Pred. No.: 6.89e-10 Length: 572
Score: 21.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.20% Indels: 0
DB: 3 Gaps: 0

US-09-993-808B-2 (1-256) x AAN02390 (1-572)
QY 225 AlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyValProLeuAspAla 244
Db 286 GCCAAACGCTTTGCTTCCAACTCGACTTCGTCGGCGGTGCCCTCGACGCC 345
QY 245 Gly 245
Db 346 GGC 348

RESULT 5
AAA95276
ID AAA95276 standard; cDNA; 572 BP.
XX
AC AAA95276;
XX
DT 17-JAN-2001 (first entry)
XX
DE Corn cyclin-dependent kinase inhibitor coding sequence #1.
XX
KW Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
KW cell growth; herbicide; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 208..366
FT /*tag= a
FT /product= "CDKI"
FT /partial
XX
PN WO200060087-A2.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US009106.
XX
PR 07-APR-1999; 99US-0128192P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX Klein TM, Weng Z, Cahoon RE;
XX

DR WPI; 2000-679375/66.
DR P-PSDB; AAB26245.
XX
PT Cyclin dependent kinase inhibitor sequences, useful for identifying
PT herbicides and plant growth inhibitors.
XX
PS Claim 2; Page 39-40; 58pp; English.
XX
CC The present sequence is the coding sequence for the corn cyclin-dependent
CC kinase inhibitor (CDKI). It was isolated by searching a corn silk cDNA
CC library for sequences similar to those encoding the CDKI from Chenopodium
CC rubrum, Caenorhabditis elegans and Arabidopsis thaliana. CDKI is involved
CC in the cell cycle, and may promote or inhibit cell division and growth.
CC The coding sequence and the protein it encodes are useful in the
CC production of transgenic plants which produce increased or decreased
CC amounts of the CDKI protein, in the identification of herbicides, in
CC genetic and physical mapping and in the isolation of the CDKI gene in
CC other organisms
XX
SQ Sequence 572 BP; 140 A; 156 C; 176 G; 93 T; 0 U; 7 Other;

Alignment Scores:
Pred. No.: 6.89e-10 Length: 572
Score: 21.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.20% Indels: 0
DB: 3 Gaps: 0

US-09-993-808B-2 (1-256) x AAA95276 (1-572)
QY 225 AlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyValProLeuAspAla 244
Db 286 GCCAAACGCTTTGCTTCCAACTCGACTTCGTCGGCGGTGCCCTCGACGCC 345
QY 245 Gly 245
Db 346 GGC 348

RESULT 6
AAA95280
ID AAA95280 standard; cDNA; 639 BP.
XX
AC AAA95280;
XX
DT 17-JAN-2001 (first entry)
XX
DE Corn cyclin-dependent kinase inhibitor coding sequence #2.
XX
KW Corn; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
KW cell growth; herbicide; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 8..388
FT /*tag= a
FT /product= "CDKI"
FT /partial
XX
PN WO200060087-A2.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US009106.
XX
PR 07-APR-1999; 99US-0128192P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX Klein TM, Weng Z, Cahoon RE;
XX WPI; 2000-679375/66.
DR

DR P-PSDB; AAB26249.

XX Cyclin dependent kinase inhibitor sequences, useful for identifying

PT herbicides and plant growth inhibitors.

XX

PS Claim 2; Page 43; 58pp; English.

XX

CC The present sequence is the coding sequence for the corn cyclin-dependent

CC kinase inhibitor (CDKI). It was isolated by searching a corn silk cDNA

CC library for sequences similar to those encoding the CDKI from Chenopodium

CC rubrum, Caenorhabditis elegans and Arabidopsis thaliana. CDKI is involved

CC in the cell cycle, and may promote or inhibit cell division and growth.

CC The coding sequence and the protein it encodes are useful in the

CC production of transgenic plants which produce increased or decreased

CC amounts of the CDKI protein, in the identification of herbicides, in

CC genetic and physical mapping and in the isolation of the CDKI gene in

CC other organisms

XX

SQ Sequence 639 BP; 169 A; 169 C; 192 G; 109 T; 0 U; 0 Other;

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 7.64e-10 | Length: | 639 |
| Score: | 21.00 | Matches: | 21 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 8.20% | Indels: | 0 |
| DB: | 3 | Gaps: | 0 |

US-09-993-808B-2 (1-256) x AAA95280 (1-639)

QY 225 AlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyValProLeuAspAla 244

Db 293 GCCAAACGCTTTCCTCCAAAGTACAACTTCGACTTCGTCGGGGCGTCCCTCGACGCC 352

QY 245 Gly 245

Db 353 GGC 355

RESULT 7

ABV74605

ID ABV74605 standard; DNA; 841 BP.

XX

AC ABV74605;

XX

DT 21-FEB-2003 (first entry)

XX

DE Maize CKI_D coding sequence.

XX

KW Maize; cyclin-dependent kinase inhibitor; CKI; CKI_D; plant; corn;

KW crop yield; root size; plant growth; tassel size; ear size;

KW male sterility; endoreduplication; gene; ds.

XX

OS Zea mays.

XX

FH Key Location/Qualifiers

FT CDS 159..839

FT /*tag= a

FT /partial

FT /product= "CKI_D"

FT /note= "No stop codon given"

XX

PN WO200281623-A2.

XX

PD 17-OCT-2002.

XX

PF 06-NOV-2001; 2001WO-US044038.

XX

PR 07-NOV-2000; 2000US-0246349P.

XX

PA (PION-) PIONEER HI-BRED INT INC.

PA (ARIZ-) ARIZONA BOARD OF REGENTS.

XX

PI Gordon-Kamm WJ, Lowe KS, Larkins BA, Dilkes BR, Sun Y;

XX

DR WPI; 2003-058511/05.

DR P-PSDB; ABB98759.

XX

PT Novel cyclin-dependent kinase polynucleotides and their encoded proteins,

PT involved in cell cycle regulation, and useful for altering cell cycle

PT protein content, cell cycle progression, cell number and composition of

PT plants.

XX

PS Claim 1; Page 68-69; 69pp; English.

XX

CC The present sequence is the coding sequence (I) for maize cyclin-

CC dependent kinase inhibitor (CKI), CKI D. (I) is useful for modulating the

CC activity of cyclin-dependent kinase (CDK) in a plant such as a corn,

CC soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice,

CC barley, oil-seed Brassica and millet. Modulating the activity of CDK,

CC preferably modulating downward is useful for providing differential

CC growth in a plant, especially a positive growth advantage and modulating

CC CDK activity upward is useful for increasing crop yield, root size, plant

CC growth, tassel size and/or ear size. Modulating CDK activity is also

CC useful for conferring male sterility and for improving transformation

CC frequencies by increasing the number of cells in cell division. CDK

CC activity can also be modulated for modulating endoreduplication in the

CC endosperm of corn, sorghum, wheat, rice, barley, and millet, where the

CC promoter used is an endosperm-preferred promoter. The cell numbers are

CC modulated in one or more tissues of a plant, comprising root, seed,

CC tassel, ear, silk, stalk, embryo, flower, grain, germ, head, leaves,

CC stem, seed, trunk, meristem or fruit. The cells are nucleus, endosperm,

CC pericarp, meristematic or leaf cells. (I) is also useful for identifying

CC maize CKI interacting proteins, by adducting the nucleic acid sequence to

CC a second nucleic acid sequence encoding a DNA-binding domain

XX

SQ Sequence 841 BP; 142 A; 247 C; 342 G; 109 T; 0 U; 1 Other;

Alignment Scores:

| | | | |
|------------------------|---------|---------------|-----|
| Pred. NO.: | 4.8e-06 | Length: | 841 |
| Score: | 17.00 | Matches: | 17 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 6.64% | Indels: | 0 |
| DB: | 7 | Gaps: | 0 |

US-09-993-808B-2 (1-256) x ABV74605 (1-841)

QY 23 GlnValValGlyValArgThrArgSerArgSerAlaAlaThrGlyGly 39

Db 225 CAGTCTCGTCGGCGTCCGGACGAGGTCCAGGTCCGGCGGGCGGCGGC 275

RESULT 8

ABK93958

ID ABK93958 standard; DNA; 1242 BP.

XX

AC ABK93958;

XX

DT 27-AUG-2002 (first entry)

XX

DE DNA encoding OsICK2 protein.

XX

KW Plant; ds; gene; inhibitor of cyclin dependent kinase; ICK.

XX

OS Oryza sativa.

XX

PN WO200228893-A2.

XX

PD 11-APR-2002.

XX

PF 29-JUN-2001; 2001WO-IB001492.

XX

PR 14-JUL-2000; 2000US-0218471P.

PR 13-OCT-2000; 2000US-0241219P.

XX

PA (CROP-) CROPDESIGN NV.

XX

PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
 PI Hatzfeld Y;
 XX
 DR WPI; 2002-471311/50.
 DR P-PSDB; ABG65670.
 XX
 PT Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
 PT to screen substrates, drugs or compounds which modulate ICK activity and
 PT treat disorders characterized by an insufficient or excessive production
 PT of ICK inhibitors.
 XX
 PS Claim 46; Fig 1; 141pp; English.
 XX

CC This invention relates to the DNA and protein sequences of novel isolated
 CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
 CC the invention may be used for treating disorders characterised by
 CC insufficient or excessive production of an ICK inhibitor. The protein of
 CC the invention may also be used to screen for naturally-occurring ICK
 CC substrates, drugs or compounds which modulate ICK activity, as well as to
 CC treat disorders characterised by insufficient or excessive production of
 CC ICK protein, forms which have decreased or aberrant activity compared to
 CC ICK wild type protein. The present sequence represents a cDNA molecule
 CC encoding an inhibitor of cyclin dependent kinase (ICK) protein of the
 CC invention
 XX
 SQ Sequence 1242 BP; 286 A; 319 C; 417 G; 220 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 6.9e-06 Length: 1242
 Score: 17.00 Matches: 17
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.64% Indels: 0
 DB: 6 Gaps: 0

US-09-993-808B-2 (1-256) x ABK93958 (1-1242)

QY 20 GluValThrGlnValValGlyValArgThrArgSerArgSerAlaAla 36
 |||||
 Db 75 GAGGTACAGCAGGTGGTGGCGTCCGGACGAGGTGAGGTCCGAGCGCG 125

RESULT 9

ABK93981
 ID ABK93981 standard; cDNA; 681 BP.
 XX
 AC ABK93981;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE Full length Rice OsICK5 cDNA sequence.
 XX
 KW Plant; ds; gene; inhibitor of cyclin dependent kinase; ICK.
 XX
 OS Oryza sativa.
 XX
 PN WO200228893-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 29-JUN-2001; 2001WO-IB001492.
 XX
 PR 14-JUL-2000; 2000US-0218471P.
 PR 13-OCT-2000; 2000US-0241219P.
 XX
 PA (CROP-) CROPDESIGN NV.
 XX
 PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
 PI Hatzfeld Y;
 XX
 DR WPI; 2002-471311/50.
 DR P-PSDB; ABG65692.
 XX
 PT Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used

PT to screen substrates, drugs or compounds which modulate ICK activity and
 PT treat disorders characterized by an insufficient or excessive production
 PT of ICK inhibitors.
 XX
 PS Claim 46; Page 139; 141pp; English.
 XX

CC This invention relates to the DNA and protein sequences of novel isolated
 CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of
 CC the invention may be used for treating disorders characterised by
 CC insufficient or excessive production of an ICK inhibitor. The protein of
 CC the invention may also be used to screen for naturally-occurring ICK
 CC substrates, drugs or compounds which modulate ICK activity, as well as to
 CC treat disorders characterised by insufficient or excessive production of
 CC ICK protein, forms which have decreased or aberrant activity compared to
 CC ICK wild type protein. The present sequence represents a cDNA molecule
 CC encoding an inhibitor of cyclin dependent kinase (ICK) protein of the
 CC invention
 XX
 SQ Sequence 681 BP; 102 A; 176 C; 307 G; 94 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.: 11.1 Length: 681
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.91% Indels: 0
 DB: 6 Gaps: 0

US-09-993-808B-2 (1-256) x ABK93981 (1-681)

QY 237 ValArgGlyValProLeuAspAlaGlyGly 246
 |||||
 Db 613 GTGCGGGCGGTGCGCTCGACGCGCGGC 642

RESULT 10

ABK93982
 ID ABK93982 standard; DNA; 1073 BP.
 XX
 AC ABK93982;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE Rice OsICK5 genomic DNA clone.
 XX
 KW Plant; ds; gene; inhibitor of cyclin dependent kinase; ICK.
 XX
 OS Oryza sativa.
 XX
 PN WO200228893-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 29-JUN-2001; 2001WO-IB001492.
 XX
 PR 14-JUL-2000; 2000US-0218471P.
 PR 13-OCT-2000; 2000US-0241219P.
 XX
 PA (CROP-) CROPDESIGN NV.
 XX
 PI Frankard VMS, Peres Bota AM, Droual A, Mironov V, Inze D;
 PI Hatzfeld Y;
 XX
 DR WPI; 2002-471311/50.
 XX
 PT Novel plant ICK (Inhibitors of Cyclin Dependent Kinases) polypeptide used
 PT to screen substrates, drugs or compounds which modulate ICK activity and
 PT treat disorders characterized by an insufficient or excessive production
 PT of ICK inhibitors.
 XX
 PS Claim 46; Page 140; 141pp; English.
 XX
 CC This invention relates to the DNA and protein sequences of novel isolated
 CC ICK (Inhibitors of Cyclin Dependent Kinases) proteins. The sequences of

CC the invention may be used for treating disorders characterised by
CC insufficient or excessive production of an ICK inhibitor. The protein of
CC the invention may also be used to screen for naturally-occurring ICK
CC substrates, drugs or compounds which modulate ICK activity, as well as to
CC treat disorders characterised by insufficient or excessive production of
CC ICK protein, forms which have decreased or aberrant activity compared to
CC ICK wild type protein. The present sequence represents a cDNA molecule
CC encoding an inhibitor of cyclin dependent kinase (ICK) protein of the
CC invention
XX
SQ Sequence 1073 BP; 172 A; 287 C; 417 G; 197 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 17 Length: 1073
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 6 Gaps: 0

US-09-993-808B-2 (1-256) x ABX93982 (1-1073)
QY 237 ValArgGlyValProLeuAspAlaGlyGly 246
Db 1005 GTGCGGGCGTGGCGCTCGACGCGCGGC 1034

RESULT 11
ACA42627
ID ACA42627 standard; DNA; 1539 BP.

XX ACA42627;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #24284.

KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.

OS Pseudomonas aeruginosa.

XX WO200277183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US0009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

DR WPI; 2003-029926/02.
DR P-PSDB; ABU38757.

XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 30497; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 1539 BP; 197 A; 517 C; 477 G; 348 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 23.8 Length: 1539
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 7 Gaps: 0

US-09-993-808B-2 (1-256) x ACA42627 (1-1539)

QY 113 LeuSerArgCysSerSerThrAlaSerSer 122

Db 998 CTGTCTCGCTGTTCAGTACGGCAAGTTCA 1027

RESULT 12

ADA71170

ID ADA71170 standard; DNA; 3525 BP.

XX ADA71170;

XX 20-NOV-2003 (first entry)

XX Rice gene, SEQ ID 4493.

XX Plant; bacterial infection; fungal infection; viral infection; rice;
KW gene; ds.

XX Oryza sativa.

XX WO2003000898-A1.

XX 03-JAN-2003.

XX 22-JUN-2001; 2001WO-IB001105.

XX 22-JUN-2001; 2001WO-IB001105.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175290/17.

PT Identifying at least one gene involved in plant resistance or response to

PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.

PS Claim 6; SEQ ID NO 4493; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.

SQ Sequence 3525 BP; 777 A; 975 C; 953 G; 814 T; 0 U; 6 Other;

Alignment Scores: Length: 3525
Pred. No.: 51.6 Matches: 10
Score: 10.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 3.91% Gaps: 0
DB: 7

US-09-993-808B-2 (1-256) x ADA71170 (1-3525)

QY 28 ArgThrArgSerArgSerAlaAlaAlaThr 37

Db 206 CGTACGCGCTCGCGCTCTGCGCGGCGACG 235

RESULT 13

AAA95286

ID AAA95286 standard; cDNA; 204 BP.

XX
AC AAA95286;

XX 17-JAN-2001 (first entry)

XX Rice cyclin-dependent kinase inhibitor coding sequence #3.

KW Rice; cyclin-dependent kinase inhibitor; CDKI; cell cycle; cell division;
KW cell growth; herbicide; ss.

XX Oryza sativa.

XX Key Location/Qualifiers
FT CDS 11..125

FT /*tag= a

FT /product= "CDKI"

FT /transl_except= (pos:110..113,aa:Ser)

FT /partial

XX WO200060087-A2.

PN 12-OCT-2000.

XX 06-APR-2000; 2000WO-US009106.

XX 07-APR-1999; 99US-0128192P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Klein TM, Weng Z, Cahoon RE;

XX WPI; 2000-679375/66.

XX P-PSDB; AAB27252.

XX Cyclin dependent kinase inhibitor sequences, useful for identifying
PT herbicides and plant growth inhibitors.

XX

PS Claim 2; Page 48; 58pp; English.

XX The present sequence is the coding sequence for the rice cyclin-dependent
CC kinase inhibitor (CDKI). It was isolated by searching a rice seed cDNA
CC library for sequences similar to those encoding the CDKI from *Chenopodium*
CC *rubrum*, *Caenorhabditis elegans* and *Arabidopsis thaliana*. CDKI is involved
CC in the cell cycle, and may promote or inhibit cell division and growth.
CC The coding sequence and the protein it encodes are useful in the
CC production of transgenic plants which produce increased or decreased
CC amounts of the CDKI protein, in the identification of herbicides, in
CC genetic and physical mapping and in the isolation of the CDKI gene in
CC other organisms

SQ Sequence 204 BP; 42 A; 49 C; 75 G; 29 T; 0 U; 9 Other;

Alignment Scores: Length: 204
Pred. No.: 30.3 Matches: 9
Score: 9.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 3.52% Gaps: 0
DB: 3

US-09-993-808B-2 (1-256) x AAA95286 (1-204)

QY 237 ValArgGlyValProLeuAspAlaGly 245

Db 59 GTTCGGCGGTGCCCCCTCGACGCCGGT 85

RESULT 14

AAN02400

ID AAN02400 standard; DNA; 204 BP.

XX
AC AAN02400;

XX 01-NOV-2001 (first entry)

XX Rice Cyclin dependent kinase inhibitor (CDKI) clone rds2c.pk008.o24.

KW Cyclin dependent kinase inhibitor; CDKI; herbicide; cell cycle; rice;
KW plant growth inhibitor; ds.

XX Oryza sativa.

XX Key Location/Qualifiers
FT CDS 11..125

FT /*tag= a

FT /product= "CDKI fragment"

FT /partial

FT /note= "No start codon given"

FT /transl_except= (pos:107..110,aa:Val)

FT /note= "This codon has an apparent 1 nucleotide insertion
which alters the reading frame"

XX WO200060087-A2.

PN 12-OCT-2000.

XX 06-APR-2000; 2000WO-US009106.

XX 07-APR-1999; 99US-0128192P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Klein TM, Weng Z, Cahoon RE;

XX WPI; 2000-679375/66.

XX P-PSDB; AAP01950.

XX Cyclin dependent kinase inhibitor sequences, useful for identifying
PT herbicides and plant growth inhibitors.

XX Claim 2; Page 48; 58pp; English.

PS

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2004, 12:56:31 ; Search time 85 Seconds
(without alignments)
1671.382 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 256
Sequence: 1 MGKYMCKRCGAAGAEVAE.....VRGVPLDAGGRFEWAPVVS 256

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Word size: 1
Total number of hits satisfying chosen parameters: 1360453

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US0993808/runat_01102004_171727_13734/app_query.fasta_1.455
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=oligo.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0993808@cgn_1_1_105@runat_01102004_171727_13734 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|---------------------|-------------------|
| 1 | 9 | 3.5 | 436 | US-09-302-769-30 | Sequence 30, Appl |
| 2 | 9 | 3.5 | 642 | US-09-252-991A-6820 | Sequence 6820, Ap |
| 3 | 9 | 3.5 | 714 | US-09-252-991A-6929 | Sequence 6929, Ap |
| 4 | 9 | 3.5 | 723 | US-09-252-991A-4826 | Sequence 4826, Ap |
| 5 | 9 | 3.5 | 771 | US-09-489-039A-4276 | Sequence 4276, Ap |
| 6 | 9 | 3.5 | 792 | US-09-489-039A-6808 | Sequence 6808, Ap |
| 7 | 9 | 3.5 | 1053 | US-09-252-991A-6360 | Sequence 6360, Ap |
| 8 | 9 | 3.5 | 1548 | US-09-252-991A-4907 | Sequence 4907, Ap |
| 9 | 9 | 3.5 | 1602 | US-09-252-991A-4836 | Sequence 4836, Ap |
| 10 | 9 | 3.5 | 1623 | US-09-252-991A-6190 | Sequence 6190, Ap |
| 11 | 9 | 3.5 | 2052 | US-09-252-991A-6112 | Sequence 6112, Ap |
| 12 | 9 | 3.5 | 34279 | US-09-596-002-26 | Sequence 26, Appl |

| | | | | | | | |
|---|----|---|-----|------|---|----------------------|-------------------|
| C | 13 | 8 | 3.1 | 372 | 4 | US-09-134-000C-2276 | Sequence 2276, Ap |
| | 14 | 8 | 3.1 | 378 | 4 | US-09-295-996B-3 | Sequence 3, Appli |
| | 15 | 8 | 3.1 | 378 | 4 | US-09-295-846B-3 | Sequence 3, Appli |
| | 16 | 8 | 3.1 | 378 | 4 | US-09-551-737C-3 | Sequence 3, Appli |
| | 17 | 8 | 3.1 | 378 | 4 | US-09-551-738B-3 | Sequence 3, Appli |
| C | 18 | 8 | 3.1 | 422 | 3 | US-09-060-756-568 | Sequence 568, App |
| C | 19 | 8 | 3.1 | 422 | 4 | US-09-670-314-568 | Sequence 568, App |
| C | 20 | 8 | 3.1 | 447 | 4 | US-09-252-991A-10231 | Sequence 10231, A |
| C | 21 | 8 | 3.1 | 471 | 2 | US-08-883-070-2 | Sequence 2, Appli |
| C | 22 | 8 | 3.1 | 471 | 4 | US-09-621-976-847 | Sequence 847, App |
| C | 23 | 8 | 3.1 | 483 | 4 | US-09-252-991A-7028 | Sequence 7028, Ap |
| C | 24 | 8 | 3.1 | 520 | 4 | US-09-621-976-3031 | Sequence 3031, Ap |
| C | 25 | 8 | 3.1 | 540 | 4 | US-09-252-991A-8395 | Sequence 8395, Ap |
| C | 26 | 8 | 3.1 | 654 | 4 | US-09-252-991A-14070 | Sequence 14070, A |
| | 27 | 8 | 3.1 | 683 | 4 | US-09-685-166A-896 | Sequence 896, App |
| C | 28 | 8 | 3.1 | 705 | 1 | US-10-095-946-17 | Sequence 17, Appl |
| C | 29 | 8 | 3.1 | 705 | 4 | US-09-183-959-17 | Sequence 17, Appl |
| C | 30 | 8 | 3.1 | 705 | 4 | US-09-535-315-17 | Sequence 17, Appl |
| C | 31 | 8 | 3.1 | 750 | 1 | US-10-095-946-18 | Sequence 18, Appl |
| C | 32 | 8 | 3.1 | 750 | 4 | US-09-183-959-18 | Sequence 18, Appl |
| C | 33 | 8 | 3.1 | 750 | 4 | US-09-535-315-18 | Sequence 18, Appl |
| | 34 | 8 | 3.1 | 764 | 4 | US-08-426-630-21 | Sequence 21, Appl |
| C | 35 | 8 | 3.1 | 765 | 4 | US-09-252-991A-1872 | Sequence 1872, Ap |
| C | 36 | 8 | 3.1 | 846 | 4 | US-09-252-991A-8327 | Sequence 8327, Ap |
| | 37 | 8 | 3.1 | 1001 | 4 | US-09-016-434-963 | Sequence 963, App |
| | 38 | 8 | 3.1 | 1017 | 4 | US-09-252-991A-10799 | Sequence 10799, A |
| C | 39 | 8 | 3.1 | 1040 | 1 | US-10-095-946-11 | Sequence 11, Appl |
| C | 40 | 8 | 3.1 | 1040 | 4 | US-09-183-959-11 | Sequence 11, Appl |
| C | 41 | 8 | 3.1 | 1040 | 4 | US-09-535-315-11 | Sequence 11, Appl |
| | 42 | 8 | 3.1 | 1077 | 4 | US-09-252-991A-10719 | Sequence 10719, A |
| C | 43 | 8 | 3.1 | 1101 | 4 | US-09-252-991A-10349 | Sequence 10349, A |
| C | 44 | 8 | 3.1 | 1113 | 4 | US-09-252-991A-8445 | Sequence 8445, Ap |
| C | 45 | 8 | 3.1 | 1173 | 4 | US-09-252-991A-10448 | Sequence 10448, A |

ALIGNMENTS

RESULT 1
US-09-302-769-30
; Sequence 30, Application US/09302769
; Patent No. 6323317
; GENERAL INFORMATION:
; APPLICANT: HILTON, Douglas J
; APPLICANT: ALEXANDER, Warren S
; APPLICANT: VINEY, Elizabeth M
; APPLICANT: WILLSON, Tracey A
; APPLICANT: RICHARDSON, Rachael T
; APPLICANT: STARR, Robyn
; APPLICANT: NICHOLSON, Sandra E
; APPLICANT: METCALF, Donald
; APPLICANT: NICOLA, Nicos A
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC AGENTS
; FILE REFERENCE: 109762
; CURRENT APPLICATION NUMBER: US/09/302,769
; CURRENT FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: 08/962,560
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 436
; TYPE: DNA
; ORGANISM: Human
US-09-302-769-30
Alignment Scores:
Pred. No.: 25.4
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.52%
DB: 4
Length: 436
Matches: 9
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-993-808B-2 (1-256) x US-09-302-769-30 (1-436)

QY 105 GlnGlyAlaAlaLeuAlaAlaGlyLeu 113
|||||
Db 357 CAGGGCGGGCTCTGGCTGCAGGTCTC 383

RESULT 2

US-09-252-991A-6820
; Sequence 6820, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 6820

; LENGTH: 642

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-6820

Alignment Scores:

| | | | |
|------------------------|---------|---------------|-----|
| Pred. No.: | 36 | Length: | 642 |
| Score: | 9.00 | Matches: | 9 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 3.52% | Indels: | 0 |
| DB: | 4 | Gaps: | 0 |

US-09-993-808B-2 (1-256) x US-09-252-991A-6820 (1-642)

QY 29 ThrArgSerArgSerAlaAlaAlaThr 37
|||||
Db 32 ACAAGATCGAGATCCGCCGCCGCAACG 58

RESULT 3

US-09-252-991A-6929/c

; Sequence 6929, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 6929

; LENGTH: 714

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-6929

Alignment Scores:

| | | | |
|------------------------|---------|---------------|-----|
| Pred. No.: | 39.5 | Length: | 714 |
| Score: | 9.00 | Matches: | 9 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 3.52% | Indels: | 0 |
| DB: | 4 | Gaps: | 0 |

US-09-993-808B-2 (1-256) x US-09-252-991A-6929 (1-714)

QY 29 ThrArgSerArgSerAlaAlaAlaThr 37
|||||
Db 626 ACAAGATCGAGATCCGCCGCCGCAACG 600

RESULT 4

US-09-252-991A-4826

; Sequence 4826, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 4826

; LENGTH: 723

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-4826

Alignment Scores:

| | | | |
|------------------------|---------|---------------|-----|
| Pred. No.: | 40 | Length: | 723 |
| Score: | 9.00 | Matches: | 9 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 3.52% | Indels: | 0 |
| DB: | 4 | Gaps: | 0 |

US-09-993-808B-2 (1-256) x US-09-252-991A-4826 (1-723)

QY 114 SerArgCysSerSerThrAlaSerSer 122
|||||
Db 96 TCTCGCTGTTCCAGTACGCGCAAGTTCA 122

RESULT 5

US-09-489-039A-4276

; Sequence 4276, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 4276

; LENGTH: 771

; TYPE: DNA

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-4276

Alignment Scores:

| | | | |
|------------------------|---------|---------------|-----|
| Pred. No.: | 42.4 | Length: | 771 |
| Score: | 9.00 | Matches: | 9 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 3.52% | Indels: | 0 |
| DB: | 4 | Gaps: | 0 |

US-09-993-808B-2 (1-256) x US-09-489-039A-4276 (1-771)

QY 50 AlaProAlaGlyGluProAlaAlaAla 58
|||||

Db 150 GCGCCTGCAGGAGAACCGCTGCAGCT 176

RESULT 6

US-09-489-039A-6808
; Sequence 6808, Application US/09489039A
; Patent No. 6610836

GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 6808

; LENGTH: 792

; TYPE: DNA

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-6808

Alignment Scores:

| | | | |
|------------------------|---------|---------------|-----|
| Pred. No.: | 43.4 | Length: | 792 |
| Score: | 9.00 | Matches: | 9 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 3.52% | Indels: | 0 |
| DB: | 4 | Gaps: | 0 |

US-09-993-808B-2 (1-256) x US-09-489-039A-6808 (1-792)

QY 50 AlaProAlaGlyGluProAlaAla 58

Db 150 GCACCTGCAGGAGAACCGCTGCAGCT 176

RESULT 7

US-09-252-991A-6360

; Sequence 6360, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 6360

; LENGTH: 1053

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-6360

Alignment Scores:

| | | | |
|------------------------|---------|---------------|------|
| Pred. No.: | 56 | Length: | 1053 |
| Score: | 9.00 | Matches: | 9 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 3.52% | Indels: | 0 |
| DB: | 4 | Gaps: | 0 |

US-09-993-808B-2 (1-256) x US-09-252-991A-6360 (1-1053)

QY 153 ValSerAlaAlaSerAsnSerGlySer 161

Db 413 GTGTCGGCGGCGTCAACAGCGTCC 439

RESULT 8

US-09-252-991A-4907/c

; Sequence 4907, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 4907

; LENGTH: 1548

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-4907

Alignment Scores:

| | | | |
|------------------------|---------|---------------|------|
| Pred. No.: | 79 | Length: | 1548 |
| Score: | 9.00 | Matches: | 9 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 3.52% | Indels: | 0 |
| DB: | 4 | Gaps: | 0 |

US-09-993-808B-2 (1-256) x US-09-252-991A-4907 (1-1548)

QY 114 SerArgCysSerSerThrAlaSerSer 122

Db 563 TCTCGCTGTTCCAGTACGCAAGTTCA 537

RESULT 9

US-09-252-991A-4836

; Sequence 4836, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 4836

; LENGTH: 1602

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-4836

Alignment Scores:

| | | | |
|------------------------|---------|---------------|------|
| Pred. No.: | 81.5 | Length: | 1602 |
| Score: | 9.00 | Matches: | 9 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 3.52% | Indels: | 0 |
| DB: | 4 | Gaps: | 0 |

US-09-993-808B-2 (1-256) x US-09-252-991A-4836 (1-1602)

QY 114 SerArgCysSerSerThrAlaSerSer 122

Db 1064 TCTCGCTGTTCCAGTACGCAAGTTCA 1090

RESULT 10

US-09-252-991A-6190/c

```
; Sequence 6190, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6190
; LENGTH: 1623
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6190

Alignment Scores:
Pred. No.: 82.4 Length: 1623
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.52% Indels: 0
DB: 4 Gaps: 0

US-09-993-808B-2 (1-256) x US-09-252-991A-6190 (1-1623)

QY 153 ValSerAlaAlaSerAsnSerGlySer 161
Db 1244 GTGTCGGCGGCGTCGACACAGCGGTCC 1218

RESULT 11
US-09-252-991A-6112/c
; Sequence 6112, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6112
; LENGTH: 2052
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6112

Alignment Scores:
Pred. No.: 102 Length: 2052
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.52% Indels: 0
DB: 4 Gaps: 0

US-09-993-808B-2 (1-256) x US-09-252-991A-6112 (1-2052)

QY 153 ValSerAlaAlaSerAsnSerGlySer 161
Db 1578 GTGTCGGCGGCGTCGACACAGCGGTCC 1552

RESULT 12
US-09-596-002-26/c
; Sequence 26, Application US/09596002
```

```
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 26
; LENGTH: 34279
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 26
; PUBLICATION INFORMATION:
US-09-596-002-26

Alignment Scores:
Pred. No.: 1.26e+03 Length: 34279
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.52% Indels: 0
DB: 4 Gaps: 0

US-09-993-808B-2 (1-256) x US-09-596-002-26 (1-34279)

QY 229 AlaSerLysTyrAsnPheAspPheVal 237
Db 15679 GCGTCTAAATATATTTTGATTTGTA 15653

RESULT 13
US-09-134-000C-2276/c
; Sequence 2276, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2276
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-2276

Alignment Scores:
Pred. No.: 164 Length: 372
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.12% Indels: 0
DB: 4 Gaps: 0

US-09-993-808B-2 (1-256) x US-09-134-000C-2276 (1-372)

QY 115 ArgCysSerSerThrAlaSerSer 122
Db 107 CGATGCTCTTCCACCGCATCTCC 84

RESULT 14
```

US-09-295-996B-3
; Sequence 3, Application US/09295996B
; Patent No. 6413530
; GENERAL INFORMATION:
; APPLICANT: Borovsky, Dov
; TITLE OF INVENTION: PESTICIDAL PEPTIDES
; FILE REFERENCE: UF-230
; CURRENT APPLICATION NUMBER: US/09/295,996B
; CURRENT FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Aedes aegypti
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(189)
; OTHER INFORMATION:
US-09-295-996B-3

Alignment Scores:
Pred. No.: 166 Length: 378
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.12% Indels: 0
DB: 4 Gaps: 0

US-09-993-808B-2 (1-256) x US-09-295-996B-3 (1-378)

Qy 154 SerAlaAlaSerAsnSerGlySer 161
Db 112 AGTGTGCTCCTCCAACTCGGGGTCG 135

RESULT 15

US-09-295-846B-3
; Sequence 3, Application US/09295846B
; Patent No. 6562590
; GENERAL INFORMATION:
; APPLICANT: Borovsky, Dov
; TITLE OF INVENTION: Transformed Cells Useful for the Control of Pests
; FILE REFERENCE: UF-223
; CURRENT APPLICATION NUMBER: US/09/295,846B
; CURRENT FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Aedes aegypti
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(189)
US-09-295-846B-3

Alignment Scores:
Pred. No.: 166 Length: 378
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.12% Indels: 0
DB: 4 Gaps: 0

US-09-993-808B-2 (1-256) x US-09-295-846B-3 (1-378)

Qy 154 SerAlaAlaSerAsnSerGlySer 161
Db 112 AGTGTGCTCCTCCAACTCGGGGTCG 135

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OMprotein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2004, 13:10:36 ; Search time 446 Seconds
(without alignments)
2909.886 Million cell updates/sec

Title: US-09-993-808B-2
Perfect score: 256
Sequence: 1 MGKYMRCRGAAGAEVAIVE.....VRGVPLDAGGRFEWAPVVS I 256

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3340653 seqs, 2534783454 residues

Word size: 1
Total number of hits satisfying chosen parameters: 6672345

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09993808/runat_01102004_171728_13817/app_query.fasta_1.455
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=oligo.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09993808 @CGN_1_1_723 @runat_01102004_171728_13817
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=60 -XGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Match | Length | DB ID | Description |
|------------|-------------|--------|-------|-------------|
|------------|-------------|--------|-------|-------------|

| | | | | | | | |
|---|----|-----|-------|------|----|----------------------|-------------------|
| | 1 | 256 | 100.0 | 1372 | 9 | US-09-993-308-1 | Sequence 1, Appli |
| | 2 | 256 | 100.0 | 1372 | 11 | US-09-993-808B-1 | Sequence 1, Appli |
| C | 3 | 117 | 45.7 | 985 | 13 | US-10-425-114-18379 | Sequence 18379, A |
| | 4 | 117 | 45.7 | 1111 | 13 | US-10-425-114-34947 | Sequence 34947, A |
| | 5 | 34 | 13.3 | 510 | 17 | US-10-767-701-24179 | Sequence 24179, A |
| | 6 | 21 | 8.2 | 568 | 16 | US-10-333-006-5 | Sequence 5, Appli |
| C | 7 | 17 | 6.6 | 834 | 13 | US-10-425-114-18085 | Sequence 18085, A |
| | 8 | 17 | 6.6 | 841 | 9 | US-09-993-308-5 | Sequence 5, Appli |
| | 9 | 17 | 6.6 | 841 | 11 | US-09-993-808B-5 | Sequence 5, Appli |
| | 10 | 17 | 6.6 | 1242 | 16 | US-10-333-006-9 | Sequence 9, Appli |
| C | 11 | 17 | 6.6 | 1671 | 17 | US-10-437-963-92632 | Sequence 92632, A |
| C | 12 | 11 | 4.3 | 803 | 13 | US-10-027-632-149457 | Sequence 149457, |
| C | 13 | 11 | 4.3 | 803 | 16 | US-10-027-632-149457 | Sequence 149457, |
| | 14 | 10 | 3.9 | 681 | 16 | US-10-333-006-54 | Sequence 54, Appl |
| | 15 | 10 | 3.9 | 1065 | 17 | US-10-437-963-19949 | Sequence 19949, A |
| | 16 | 10 | 3.9 | 1073 | 16 | US-10-333-006-56 | Sequence 56, Appl |
| C | 17 | 10 | 3.9 | 1398 | 17 | US-10-437-963-19960 | Sequence 19960, A |
| | 18 | 10 | 3.9 | 1539 | 13 | US-10-282-122A-30497 | Sequence 30497, A |
| C | 19 | 10 | 3.9 | 1923 | 17 | US-10-437-963-7641 | Sequence 7641, Ap |
| C | 20 | 10 | 3.9 | 2347 | 16 | US-10-108-260A-2400 | Sequence 2400, Ap |
| | 21 | 9 | 3.5 | 334 | 9 | US-09-867-701-4973 | Sequence 4973, Ap |
| C | 22 | 9 | 3.5 | 414 | 10 | US-09-918-995-17513 | Sequence 17513, A |
| | 23 | 9 | 3.5 | 436 | 9 | US-09-908-805B-30 | Sequence 30, Appl |
| | 24 | 9 | 3.5 | 484 | 13 | US-10-282-122A-22965 | Sequence 22965, A |
| | 25 | 9 | 3.5 | 493 | 16 | US-10-333-006-1 | Sequence 1, Appli |
| C | 26 | 9 | 3.5 | 500 | 15 | US-10-029-386-11417 | Sequence 11417, A |
| | 27 | 9 | 3.5 | 500 | 15 | US-10-029-386-11723 | Sequence 11723, A |
| C | 28 | 9 | 3.5 | 546 | 15 | US-10-029-386-11544 | Sequence 11544, A |
| | 29 | 9 | 3.5 | 636 | 13 | US-10-425-114-28124 | Sequence 28124, A |
| | 30 | 9 | 3.5 | 688 | 13 | US-10-425-114-28165 | Sequence 28165, A |
| | 31 | 9 | 3.5 | 710 | 13 | US-10-027-632-21992 | Sequence 21992, A |
| | 32 | 9 | 3.5 | 710 | 16 | US-10-027-632-21992 | Sequence 21992, A |
| C | 33 | 9 | 3.5 | 775 | 13 | US-10-027-632-167114 | Sequence 167114, |
| C | 34 | 9 | 3.5 | 775 | 16 | US-10-027-632-167114 | Sequence 167114, |
| | 35 | 9 | 3.5 | 846 | 13 | US-10-027-632-170305 | Sequence 170305, |
| | 36 | 9 | 3.5 | 846 | 13 | US-10-027-632-170306 | Sequence 170306, |
| | 37 | 9 | 3.5 | 846 | 16 | US-10-027-632-170305 | Sequence 170305, |
| | 38 | 9 | 3.5 | 846 | 16 | US-10-027-632-170306 | Sequence 170306, |
| C | 39 | 9 | 3.5 | 856 | 15 | US-10-029-386-25126 | Sequence 25126, A |
| | 40 | 9 | 3.5 | 974 | 15 | US-10-029-386-22605 | Sequence 22605, A |
| | 41 | 9 | 3.5 | 975 | 13 | US-10-107-431-14 | Sequence 14, Appl |
| | 42 | 9 | 3.5 | 1083 | 16 | US-10-260-238-271 | Sequence 271, App |
| C | 43 | 9 | 3.5 | 1264 | 16 | US-10-264-049-116 | Sequence 116, App |
| C | 44 | 9 | 3.5 | 1392 | 13 | US-10-282-122A-31290 | Sequence 31290, A |
| C | 45 | 9 | 3.5 | 1395 | 13 | US-10-282-122A-30635 | Sequence 30635, A |

ALIGNMENTS

RESULT 1
US-09-993-308-1
; Sequence 1, Application US/099933308
; Patent No. US20020159435A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William J.
; APPLICANT: Lowe, Keith S.
; APPLICANT: Larkins, Brian A.
; APPLICANT: Dikes, Brian R.
; APPLICANT: Sun, Yuejin
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,308
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 60/246,349
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1372
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:

; NAME/KEY: CDS
; LOCATION: (134)..(904)
; OTHER INFORMATION:
US-09-993-308-1

Alignment Scores:

Pred. No.: 1.49e-233 Length: 1372
Score: 256.00 Matches: 256
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-993-808B-2 (1-256) x US-09-993-308-1 (1-1372)

```
QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
Db 134 ATGGGAAGTACATCGCAAGTCCAGGGCGCGCGAGGTCCGCCGCGAGGTCCGCCGCGAG 193
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyVal 40
Db 194 GTTACGCAGGTCGTCCGCGTCCGACGAGGTCCAGGTCCCGCGCGCGAGGTCCGCCGCGAG 253
QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaAlaValSer 60
Db 254 GCGAAGTCCGCCCGAGGAGGAGAGGGCGCGCGCGCGCGAGGTCCGCCGCGAGGTCCGCCGCGAG 313
QY 61 AlaGlyGlyAspGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAla 80
Db 314 GCTGGTGGGACGGCGGAGGTCTACATCCACCTGCGTAGCCGATGCTGTTCATGGCA 373
QY 81 ProProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAspGly 100
Db 374 CCGCTCAGCCCGAGCGCGCTCGGTTCGGTCCGACCCCGGTGGAGGCTGCTGATGGC 433
QY 101 AlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThrAla 120
Db 434 GCTCAGGACAGCAGCGCGCGCTCGCGCGCGGCTCGCGTTGCTCCAGCACGGCG 493
QY 121 SerSerValAsnLeuGlyLeuGlyGlnArgGlySerHisThrCysArgSerTyrAsp 140
Db 494 TCGTCGGTGAACTTGGGCTTGGGGGTTCAGCGCGGAGCACACCTGCCGCTCCTACGAC 553
QY 141 AlaAlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGly 160
Db 554 GCTGCAGAGGCTGGCGGGATCACGTCCTGGTGTGATGTCGGCGCGGAGCAACTCCGGG 613
QY 161 SerGlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyGluLeu 180
Db 614 AGCGCCCGAGACCGCGAGAGGCGAGAGACGACGACATCCAGCCCGGCGCACGGCGAGCTC 673
QY 181 SerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThr 200
Db 674 AGCGATCTGGAGTCGATCGTGGCGGGCACAAAGACTGGCCCGTCCGTAACCGCGGCAACG 733
QY 201 ProAlaAlaGluLeuValProProAlaHisGluLeuGlnPhePheAlaAlaAla 220
Db 734 CCGCTCGGAGCTGATCGTCCGCGCAGCACACGAGATCCAGGAGTTCTTCGCGCGCGCC 793
QY 221 GluAlaAlaGlnAlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyVal 240
Db 794 GAGCGGCCCGAGGCCAAGCGTTTGTCTTCCAAGTACAACTTCGACTTCGTCGCGCGCGTG 853
QY 241 ProLeuAspAlaGlyGlyArgPheGluTrpAlaProValValSerIle 256
Db 854 CCCCTCGACCGCGCGCGGTCGAGTGGCGCGCGCGGTCGAGTCCGTCAGCATC 901
```

RESULT 2

US-09-993-808B-1

; Sequence 1, Application US/09993808B

; Publication No. US2004000343A1

; GENERAL INFORMATION:

; APPLICANT: Gordon-Kamm, William

; APPLICANT: Lowe, Keith
; APPLICANT: Sun, Yuejin
; APPLICANT: Dilkes, Brian
; APPLICANT: Larkins, Brian
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides,
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,808B
; CURRENT FILING DATE: 2001-11-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1372
; TYPE: DNA
; ORGANISM: zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (134)..(902)
US-09-993-808B-1

Alignment Scores:

Pred. No.: 1.49e-233 Length: 1372
Score: 256.00 Matches: 256
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0

US-09-993-808B-2 (1-256) x US-09-993-808B-1 (1-1372)

```
QY 1 MetGlyLysTyrMetArgLysCysArgGlyAlaAlaGlyAlaGluValAlaAlaValGlu 20
Db 134 ATGGGAAGTACATCGCAAGTCCAGGGCGCGCGAGGTCCGCCGCGAGGTCCGCCGCGAG 193
QY 21 ValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyVal 40
Db 194 GTTACGCAGGTCGTCCGCGTCCGACGAGGTCCAGGTCCCGCGCGGAGGTCCGCCGCGAG 253
QY 41 AlaLysValAlaProArgArgLysArgAlaProAlaGlyGluProAlaAlaAlaValSer 60
Db 254 GCGAAGTCCGCCCGAGGAGGAGGCGCGCGCGCGCGAGGTCCGCCGCGGAGCTGCTGCCGCCGAG 313
QY 61 AlaGlyGlyAspGlyGlySerCysTyrIleHisLeuArgSerArgMetLeuPheMetAla 80
Db 314 GCTGGTGGGACGGCGGAGGTCTACATCCACCTGCGTAGCCGATGCTGTTCATGGCA 373
QY 81 ProProGlnProGlnProSerValAspSerValProThrProValGluAlaAlaAspGly 100
Db 374 CCGCTCAGCCCGAGCGCGTTCGTTGACTCGGTTCCGACCCCGGTGGAGGCTGCTGATGGC 433
QY 101 AlaAlaGlyGlnGlnGlyAlaAlaLeuAlaAlaGlyLeuSerArgCysSerSerThrAla 120
Db 434 GCTCAGGACAGCAGCGCGCGCGCTCGCGCGCGGCTTCGCGTTGCTCCAGCACGGCG 493
QY 121 SerSerValAsnLeuGlyLeuGlyGlnArgGlySerHisThrCysArgSerTyrAsp 140
Db 494 TCGTCGGTGAACTTGGGCTTGGGGGTTCAGCGCGGAGCACACCTGCCGCTCCTACGAC 553
QY 141 AlaAlaGluAlaGlyGlyAspHisValLeuValAspValSerAlaAlaSerAsnSerGly 160
Db 554 GCTGCAGAGGCTGGCGGGATCACGTCCTGGTGTGATGTCGGCGCGGAGCAACTCCGGG 613
QY 161 SerGlyProAspArgGluArgGluThrThrProSerSerArgAlaHisGlyGluLeu 180
Db 614 AGCGCCCGAGACCGCGAGGCGAGAGACGACGACATCCAGGAGTTCTTCGCGCGCGCC 673
QY 181 SerAspLeuGluSerAspLeuAlaGlyHisLysThrGlyProSerLeuProAlaAlaThr 200
Db 674 AGCGATCTGGAGTCGATCGTGGCGGGCACAAAGACTGGCCCGTCCGTAACCGCGGCAACG 733
QY 201 ProAlaAlaGluLeuValProProAlaHisGluLeuGlnPhePheAlaAlaAla 220
Db 734 CCGCTCGGAGCTGATCGTCCGCGCAGCACACGAGATCCAGGAGTTCTTCGCGCGCGCC 793
```


OTHER INFORMATION: Clone ID: 30165606
US-10-767-701-24179

Alignment Scores:

Pred. No.: 5.54e-23 Length: 510
Score: 34.00 Matches: 34
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 13.28% Indels: 0
DB: 17 Gaps: 0

US-09-993-808B-2 (1-256) x US-10-767-701-24179 (1-510)

QY 15 GluValAlaAlaValGluValThrGlnValValGlyValArgThrArgSerArgSerAla 34
Db 166 GAGGTGCGCCGAGTCGAGTCACGAGGTGTCTCGGCTCAGGACGAGGTCCAGGTCGCG 225
QY 35 AlaAlaThrGlyGlyValAlaLysValAlaProArgArgLys 48
Db 226 GCGGCGACCGGGGTGTCTCGGAAGGTCTCCGAGGAGGAAG 267

RESULT 6

US-10-333-006-5
Sequence 5, Application US/103333006
Publication No. US20040019926A1
GENERAL INFORMATION:
APPLICANT: Frankard, Valerie Marie-No. US20040019926A1lle S.
APPLICANT: Peres Bota, Adrian Marius
APPLICANT: Droual, Anne-Marie
APPLICANT: Mironov, Vladimir
APPLICANT: Inz, Dirk
APPLICANT: Hatzfeld, Yves

TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS

FILE REFERENCE: 1187-13
CURRENT APPLICATION NUMBER: US/10/333,006
CURRENT FILING DATE: 2003-01-14
PRIOR APPLICATION NUMBER: PCT/IB01/01492
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: US 60/218,471
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/241,219
PRIOR FILING DATE: 2000-10-13
NUMBER OF SEQ ID NOS: 59
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 568
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc feature
LOCATION: (32)..(32)
OTHER INFORMATION: n = A, T, C or G

Alignment Scores:

Pred. No.: 1.36e-10 Length: 568
Score: 21.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.20% Indels: 0
DB: 16 Gaps: 0

US-09-993-808B-2 (1-256) x US-10-333-006-5 (1-568)

QY 225 AlaLysArgPheAlaSerLysTyrAsnPheAspPheValArgGlyValProLeuAspAla 244
Db 257 GCCAAACGCTTTGCTTCCAAGTACAACTTCGACTTCGTCGCGGGGTGCCCTCGACGCC 316
QY 245 Gly 245
Db 317 GGC 319

RESULT 7

US-10-425-114-18085/c
Sequence 18085, Application US/10425114
Publication No. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53313)B

CURRENT APPLICATION NUMBER: US/10/425,114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 18085

LENGTH: 834

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Clone ID: LIB3078-022-G5_FLI

US-10-425-114-18085

Alignment Scores:

Pred. No.: 1.17e-06 Length: 834
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.64% Indels: 0
DB: 13 Gaps: 0

US-09-993-808B-2 (1-256) x US-10-425-114-18085 (1-834)

QY 23 GlnValValGlyValArgThrArgSerArgSerAlaAlaThrGlyGly 39

Db 630 CAGGTCGTCCGGCTCCGACGAGGTCCAGGTCCGCGCGGCGACCGCGGC 580

RESULT 8

US-09-993-308-5
Sequence 5, Application US/09993308
Patent No. US20020159435A1

GENERAL INFORMATION:

APPLICANT: Gordon-Kamm, William J.
APPLICANT: Lowe, Keith S.
APPLICANT: Larkins, Brian A.
APPLICANT: Dilkes, Brian R.
APPLICANT: Sun, Yuejin

TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides and Uses Thereof

FILE REFERENCE: 1146

CURRENT APPLICATION NUMBER: US/09/993,308

CURRENT FILING DATE: 2001-11-06

PRIOR APPLICATION NUMBER: 60/246,349

PRIOR FILING DATE: 2000-11-07

NUMBER OF SEQ ID NOS: 6

SOFTWARE: PatentIn version 3.1

SEQ ID NO 5

LENGTH: 841

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

NAME/KEY: CDS

LOCATION: (159)..(839)

OTHER INFORMATION: The 'r' at location 491 stands for g or a.

US-09-993-308-5

Alignment Scores:

Pred. No.: 1.18e-06 Length: 841
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.64% Indels: 0
DB: 9 Gaps: 0

```
US-09-993-808B-2 (1-256) x US-09-993-308-5 (1-841)
QY      23 GlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGly 39
Db      225 CAGGTCGTCGGCGTCCGGACGAGGTCCAGGTCGCGCGCGGACCGCGGC 275

RESULT 9
US-09-993-808B-5
; Sequence 5, Application US/09993808B
; Publication No. US20040003433A1
; GENERAL INFORMATION:
; APPLICANT: Gordon-Kamm, William
; APPLICANT: Lowe, Keith
; APPLICANT: Sun, Yuejin
; APPLICANT: Dilkes, Brian
; APPLICANT: Larkins, Brian
; TITLE OF INVENTION: Cell Cycle Nucleic Acids, Polypeptides,
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: 1146
; CURRENT APPLICATION NUMBER: US/09/993,808B
; CURRENT FILING DATE: 2001-11-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 841
; TYPE: DNA
; ORGANISM: zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (159)...(841)
; OTHER INFORMATION: r= g or a
US-09-993-808B-5

Alignment Scores:
Pred. No.:      1.18e-06      Length:      841
Score:          17.00         Matches:      17
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    6.64%        Indels: 0
DB:             11           Gaps: 0

US-09-993-808B-2 (1-256) x US-09-993-808B-5 (1-841)
QY      23 GlnValValGlyValArgThrArgSerArgSerAlaAlaAlaThrGlyGly 39
Db      225 CAGGTCGTCGGCGTCCGGACGAGGTCCAGGTCGCGCGCGGACCGCGGC 275

RESULT 10
US-10-333-006-9
; Sequence 9, Application US/10333006
; Publication No. US20040019926A1
; GENERAL INFORMATION:
; APPLICANT: Frankard, Valerie Marie-No. US20040019926A1lle S.
; APPLICANT: Peres Bota, Adrian Marius
; APPLICANT: Droual, Anne-Marie
; APPLICANT: Mironov, Vladimir
; APPLICANT: Inz,, Dirk
; APPLICANT: Hatzfeld, Yves
; TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
; FILE REFERENCE: 1187-13
; CURRENT APPLICATION NUMBER: US/10/333,006
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: PCT/IB01/01492
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/218,471
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/241,219
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1242
```

```
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-333-006-9

Alignment Scores:
Pred. No.:      1.61e-06      Length:      1242
Score:          17.00         Matches:      17
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    6.64%        Indels: 0
DB:             16           Gaps: 0

US-09-993-808B-2 (1-256) x US-10-333-006-9 (1-1242)
QY      20 GluValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAla 36
Db      75 GAGGTCACGCAGGTGGTGGCGTCCGGACGAGGTCCGAGGTCCGACGGCG 125

RESULT 11
US-10-437-963-92632/c
; Sequence 92632, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 92632
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91095C.1
US-10-437-963-92632

Alignment Scores:
Pred. No.:      2.05e-06      Length:      1671
Score:          17.00         Matches:      17
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    6.64%        Indels: 0
DB:             17           Gaps: 0

US-09-993-808B-2 (1-256) x US-10-437-963-92632 (1-1671)
QY      20 GluValThrGlnValValGlyValArgThrArgSerArgSerAlaAlaAla 36
Db      1149 GAGGTCACGCAGGTGGTGGCGTCCGGACGAGGTCCGAGGTCCGACGGCG 1099

RESULT 12
US-10-027-632-149457/c
; Sequence 149457, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
```

; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 149457
; LENGTH: 803
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-149457

Alignment Scores:
Pred. No.: 0.568 Length: 803
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.30% Indels: 0
DB: 13 Gaps: 0

US-09-993-808B-2 (1-256) x US-10-027-632-149457 (1-803)

QY 107 AlaAlaLeuAlaAlaGlyLeuSerArgCysSer 117
|||||
Db 135 GCAGCGCTGGCGCAGGACTGAGCGGTGCTCC 103

RESULT 13

US-10-027-632-149457/c
; Sequence 149457, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 10827.129

; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 149457
; LENGTH: 803
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-149457

Alignment Scores:
Pred. No.: 0.568 Length: 803
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.30% Indels: 0
DB: 16 Gaps: 0

US-09-993-808B-2 (1-256) x US-10-027-632-149457 (1-803)

QY 107 AlaAlaLeuAlaAlaGlyLeuSerArgCysSer 117
|||||
Db 135 GCAGCGCTGGCGCAGGACTGAGCGGTGCTCC 103

RESULT 14

US-10-333-006-54
; Sequence 54, Application US/10333006
; Publication No. US20040019926A1
; GENERAL INFORMATION:
; APPLICANT: Frankard, Valerie Marie-No. US20040019926A1lle S.
; APPLICANT: Peres Bota, Adrian Marius
; APPLICANT: Droual, Anne-Marie
; APPLICANT: Mironov, Vladimir
; APPLICANT: Inz,, Dirk
; APPLICANT: Hatzfeld, Yves
; TITLE OF INVENTION: NOVEL PLANT CYCLIN-DEPENDENT KINASE INHIBITORS
; FILE REFERENCE: 1187-13
; CURRENT APPLICATION NUMBER: US/10/333,006
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: PCT/IB01/01492
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/218,471
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/241,219
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: probe or primer
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (355)..(356)
; OTHER INFORMATION: n = A, T, C or G
US-10-333-006-54

Alignment Scores:
Pred. No.: 4.43 Length: 681
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 16 Gaps: 0

US-09-993-808B-2 (1-256) x US-10-333-006-54 (1-681)

QY 237 ValArgGlyValProLeuAspAlaGlyGly 246
|||||
Db 613 GTGCGGGCGGTGCCGCTCGACGCGGGCGGC 642

RESULT 15

US-10-437-963-19949
; Sequence 19949, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 19949
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_2535C.1
US-10-437-963-19949

Alignment Scores:
Pred. No.: 6.37 Length: 1065
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 17 Gaps: 0

US-09-993-808B-2 (1-256) x US-10-437-963-19949 (1-1065)

QY 237 ValArgGlyValProLeuAspAlaGlyGly 246
|||
Db 682 GTGGGGGGGTGCCGCTCGACGCGCGGC 711

Search completed: October 2, 2004, 15:47:27
Job time : 451 secs

This Page Blank (uspto)